



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/11 // 15/63</b>	<b>A1</b>	(11) International Publication Number: <b>WO 98/07846</b> (43) International Publication Date: <b>26 February 1998 (26.02.98)</b>
<p>(21) International Application Number: <b>PCT/DK97/00342</b></p> <p>(22) International Filing Date: <b>25 August 1997 (25.08.97)</b></p> <p>(30) Priority Data: <b>0886/96</b>                      <b>23 August 1996 (23.08.96)</b>                      <b>DK</b></p> <p>(71)(72) Applicant and Inventor: <b>JENSEN, Peter, Ruhdal</b> <b>[DK/DK]; Søgårdsvej 19, DK-2820 Gentofte (DK).</b></p> <p>(72) Inventor; and (75) Inventor/Applicant (for US only): <b>HAMMER, Karin</b> <b>[DK/DK]; Gærdesmuttevej 20, DK-2970 Hørsholm (DK).</b></p> <p>(74) Agent: <b>HOFMAN-BANG &amp; BOUTARD, LEHMANN &amp; REE</b> <b>A/S; Hans Bekkevolds Allé 7, DK-2900 Hellerup (DK).</b></p>	<p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</b></p> <p><b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>	
<p>(54) Title: <b>ARTIFICIAL PROMOTER LIBRARIES FOR SELECTED ORGANISMS AND PROMOTERS DERIVED FROM SUCH LIBRARIES</b></p> <p>(57) Abstract</p> <p>An artificial promoter library for a selected organism or group of organisms is constructed as a mixture of double stranded DNA fragments, the sense strands of which comprise at least two consensus sequences of efficient promoters from said organism or group of organisms, or parts thereof comprising at least half of each, and surrounding or intermediate nucleotide sequences (spacers) of variable length in which at least 7 nucleotides are selected randomly among the nucleobases A, T, C and G. The sense strands of the double stranded DNA fragments may also include a regulatory DNA sequence imparting a specific regulatory feature, such as activation by a change in the growth conditions, to the promoters of the library. Further, they may have a sequence comprising one or more recognition sites for restriction endonucleases added to one of or both their ends. The selected organism or group or organisms may be selected from prokaryotes and from eukaryotes; and in prokaryotes the consensus sequences to be retained most often will comprise the -35 signal (-35 to -30): TTGACA and the -10 signal (-12 to -7): TATAAT or parts of both comprising at least 3 conserved nucleotides of each, while in eukaryotes said consensus sequences should comprise a TATA box and at least one upstream activation sequence (UAS). Such artificial promoter libraries can be used i.a. for optimizing the expression of specific genes in various selected organisms.</p>		

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Artificial promoter libraries for selected organisms and  
promoters derived from such libraries

**FIELD OF THE INVENTION**

5 This invention concerns artificial promoter libraries and  
a method of constructing an artificial promoter library  
for a selected organism or group of organisms. The inven-  
tion also concerns the individual novel promoters derived  
10 from such libraries. Further, the invention concerns a  
method of optimizing the expression of a gene in a se-  
lected organism by use of promoters from such an artifi-  
cial promoter library for that organism. In principle,  
artificial promoter libraries according to the invention  
15 can be constructed for use in any living organism, but  
presently they have mostly been of value for modulating  
gene expression of microorganisms. In connection with  
this invention the term "microorganism" shall be taken  
broadly to include prokaryotic organisms such as bacteria  
20 as well as eukaryotic microorganisms such as yeasts,  
other fungi and cell lines of higher organisms.

**BACKGROUND OF THE INVENTION**

25 Metabolic engineering of living organisms is still in its  
infancy with respect to industrial applications, de-  
spite the fact that genetic engineering has now been fea-  
sible for more than a decade. To a large extent, this may  
be due to the disappointing outcome of many of the at-  
tempts so far to improve strain performance. There are at  
30 least two reasons for the negative outcome of the at-  
tempts to increase metabolic fluxes:

One is that the genetic engineer tends to overlook the  
35 subtlety of control and regulation of cellular metabo-  
lism. The expression of enzymes that are expected to be  
rate limiting are increased 10 to 100 fold, e.g. by

placing the gene on a high copy number plasmid. Or, a branching flux in a pathway is eliminated by deleting a gene. Quite often, this will have secondary effects on the metabolism, for instance by lowering metabolite concentrations that are essential to other parts of the cellular metabolism (e.g. processes that are essential to the growth of the organism) and the net result may be that the overall performance of the cell with respect to the desired product is decreased. Instead, it is necessary to tune the expression of the relevant gene around the normal expression level and determine the optimal expression level, for instance as the level that maximizes or minimizes the flux.

The second reason for the negative outcome lies in the rate limiting concept itself: both metabolic control theory (Kacser and Burns, 1973) and experimental determinations of control by individual steps in a pathway (Schaaff et al., 1989; Jensen et al., 1993) have shown that reaction steps which were expected to be rate limiting with respect to a particular flux, turned out to have no or very little control over the flux. Instead, the control and regulation of the cellular metabolism turned out to be distributed over several enzymes in a pathway, and it may be necessary to enhance the expression of several enzymes in order to obtain a higher flux.

According to metabolic control theory, the total flux control exerted by all the enzymes in a pathway, should always sum up to 1. Therefore, after one enzyme concentration has been optimized, the flux control will have shifted to another enzyme(s), and it may then be useful to perform additional rounds of enzyme optimization in order to increase the flux further.

In summary, flux optimization requires 1) fine-tuning of enzyme concentration rather than many fold overexpression

and often 2) optimization of the level of several enzymes in a pathway rather than looking for the rate limiting step.

5 There are now many systems available that allow one to increase the gene expression more than 1000 fold and/or to turn on gene expression at a particular time point during a fermentation process (e.g. using temperature inducible systems). With respect to tuning the steady state  
10 gene expression in the fermenter, to say 150% or 70% of the normal expression level, it becomes more difficult. In principle, one could use a *lac*-type promoter in front of the gene of interest, and then add a certain amount of an inducer of the *lac* system, for instance IPTG  
15 (isopropyl- $\beta$ -D-thiogalactoside), or use a temperature sensitive system at the correct temperature. These possibilities are often not practical for large scale industrial applications. The alternative is to use a promoter that has exactly the right strength. However, such pro-  
20 moters are seldom available, and furthermore one needs a range of promoter activities in order to optimize the expression of the gene in the first place, see below.

During the past two decades, much work has been done to  
25 define and optimize the consensus sequences of microorganisms. In many prokaryotes, one often finds two more or less conserved DNA sequences at approximately position -10 and -35 relative to the start site for transcription, TATAAT and TTGACA, respectively, with approximately 17  
30 basepair between the two. The dogma in this field is that, by including these elements, the resulting promoters would tend to become strong. Indeed, promoter up mutations, which are relatively rare events, ususally results in a better match to the above consensus sequences,  
35 while down mutations results in a poorer match to the consensus sequences or a less optimal distances between

these. In addition, when random DNA sequences are cloned in place of one of the two consensus sequences, the strength of the resulting promoters usually correlate with the degree of homology to the consensus sequences.

5

In principle, modulation of the strength of promoters could then be achieved by basepair changes in the consensus sequences or by changes in the length of the spacer between these. But the impact of such changes on the promoter strength will tend to be large (see example 1 of this invention), and it is therefore not feasible to achieve small steps of strength modulation through basepair changes in the consensus sequences.

15

While the length of the spacer separating the two consensus sequences is known to play an important role for the strength of a promoter, the sequence of the spacers between the consensus sequences has usually been considered to be of little importance for the strength of the promoters, and attempts to identify additional consensus sequences in the spacer region through mutagenesis have indeed been unsuccessful. So far, nobody has attempted to randomize the spacer, while keeping the consensus sequences and the spacer length relatively constant.

25

Numerous experiments have been carried out in order to define and optimize the consensus sequences of microorganisms, including experiments where at least one of the consensus sequences was being randomized. In some of these experiments a part of the nucleotides surrounding the consensus sequences was also randomized in order to allow for the generation of promoters with length of spacers different from 17 bp, and/or in order to find possible new consensus motifs around the consensus sequences. The chances that this will generate an efficient

35

promoter is very small and a selection must be applied in order to find those rare cases where the homology to the consensus sequence is sufficiently high to result in even a weak promoter.

5

#### OBJECT OF THE INVENTION

The promoter libraries that we are aiming at should cover the entire range of promoter activities that could become  
10 of interest for engineering of a particular species, preferably from the weakest promoters detectable, to the strongest promoters possible. Moreover, we aim at covering this broad activity range in small steps, say increase in activity by 50-100% per step in order to be  
15 suitable for the purpose of flux optimization as described above.

In this invention, we show that the sequence of the spacers between the consensus sequences are far more important than it has previously been appreciated. The spacer  
20 sequence can have a strong impact on promoter strength, when 1) a major part of the spacer sequence is varied simultaneously and in a random manner and 2), at the same time, at least half of the consensus sequences are kept  
25 constant. We show that if these two conditions are fulfilled, our method can be used to generate promoters that cover wide ranges of activities, including very strong promoters. The range of promoter activities is covered in small steps of activity change which makes these promoters  
30 very suitable for metabolic engineering purposes. In addition, we show that the method can be used to generate promoters for a wide range of organisms and at an unusual high frequency.

**SUMMARY OF THE INVENTION**

The present invention provides an artificial promoter library for a selected organism or group of organisms, comprising a mixture of double stranded DNA fragments the sense strands of which comprise at least two consensus sequences of efficient promoters from said organism or group of organisms, or parts thereof comprising at least half of each, and surrounding or intermediate nucleotide sequences (spacers) of variable length in which at least 7 nucleotides are selected randomly among the nucleobases A, T, C and G, with the proviso that previously known promoter sequences and promoter sequences isolated from natural sources are not comprised.

The broadest variation in promoter strengths is obtained when at least 10, preferably at least 12, and more preferably at least 14, nucleotides in the spacer sequence(s) are selected randomly among the nucleobases A, T, C and G.

The sense strands of the double stranded DNA fragments may also include a regulatory DNA sequence imparting a specific regulatory feature to the promoters of the library. Such specific regulatory feature is preferably activation by a change in the growth conditions, such as a change in the pH, osmolarity, temperature or growth phase.

For cloning purposes the double stranded DNA fragments usually have sequences comprising one or more recognition sites for restriction endonucleases added to one of or both their ends; most conveniently sequences specifying multiple recognition sites for restriction endonucleases (multiple cloning sites MCS).



The selected organism or group of organisms may be selected from prokaryotes and from eukaryotes, in particular from prokaryotes and eukaryotic microorganisms such as yeasts, other fungi and cell lines of higher organisms.

An interesting group of prokaryotes i.a. in the dairy industry consists of lactic acid bacteria of the genera *Lactococcus*, *Streptococcus*, *Enterococcus*, *Lactobacillus* and *Leuconostoc*, in particular strains of the species *Lactococcus lactis* and *Streptococcus thermophilus*. Other interesting prokaryotes are bacteria belonging to the genera *Escherichia*, *Bacillus* and *Pseudomonas*, in particular the species *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas putida*.

In an artificial promoter library for prokaryotes said consensus sequences may for example comprise the -10 signal (-12 to -7): TATAAT and at least one activator protein binding site upstream of the -10 signal or parts thereof comprising at least 3 conserved nucleotides of each.

Most often the consensus sequences to be retained in an artificial promoter library for prokaryotes will comprise the -35 signal (-35 to -30): TTGACA and the -10 signal (-12 to -7): TATAAT or parts of both comprising at least 3 conserved nucleotides of each.

More efficient promoters are usually obtained when said consensus sequences comprise from 4 to 6 conserved nucleotides of the -35 signal and from 4 to 6 conserved nucleotides of the -10 signal, preferably 5 or 6, and more preferably all 6 nucleotides of each. The most efficient promoters are obtained when said consensus sequences further comprise intervening conserved motifs, e.g. selected

from the conserved motifs -44 to -41: AGTT, -40 to -36: TATTC, -15 to -14: TG, and +1 to +8: GTACTGTT.

5 In such promoters the length of the spacer between the -35 signal and the -10 signal should be 14-23 bp, preferably 16-18 bp, and more preferably 17 bp. This should be understood to mean the spacer length between the hexamer signals, even when some of the nucleotides in the signals have been mutated.

10 In eukaryotic organisms said consensus sequences should comprise a TATA box and at least one upstream activation sequence (UAS).

15 An interesting eukaryotic microorganism is the yeast species *Saccharomyces cerevisiae*, normal baker's yeast. In promoters to be used in *Saccharomyces* the consensus sequences may further comprise a transcription initiation signal (TI box) functioning in *Saccharomyces cerevisiae*.

20 In a specific embodiment of an artificial promoter library according to the invention for *Saccharomyces cerevisiae* said consensus sequences comprise the TATA box: TATAAA, the UAS<sub>GCN4p</sub>: TGAACA, and the TI box:  
25 CTCTTAAGTGCAAGTGACTGCGA, which also functions as the binding site for the arginine repressor, *argR*.

The individual promoters of the artificial promoter libraries defined above are also comprised by the invention.  
30 Specific promoters which have been constructed according to the following examples are those stated in the SEQ IDs Nos. 5 to 58 below. The invention further comprises artificial promoters which are derived from promoters defined by the artificial promoter libraries of  
35 the invention.

The present invention also provides a method of constructing an artificial promoter library for a selected organism or group of organisms, which comprises selecting at least two consensus sequences of efficient promoters from said organism or group of organisms; synthesizing a mixture of single stranded DNA sequences comprising said consensus sequences, or parts thereof comprising at least half of each, and surrounding or intermediate nucleotide sequences (spacers) of variable length in which at least 7 nucleotides are selected randomly among the nucleobases A, T, C and G; and converting the single stranded DNA sequences into double stranded DNA fragments by second strand synthesis.

As previously mentioned, the broadest variation is obtained when at least 10, preferably at least 12, and more preferably at least 14, nucleotides in the spacer sequence(s) are selected randomly among the nucleobases A, T, C and G.

In order to obtain an artificial promoter library which is susceptible to regulation, the single stranded DNA sequences which are synthesized may include a regulatory DNA sequence imparting a specific regulatory feature to the promoters of the library. Such specific regulatory feature is preferably activation by a change in the growth conditions, such as a change in the pH, osmolarity, temperature or growth phase.

Also, in order to obtain an artificial promoter library suitable for cloning, a sequence specifying one or more recognition sites for restriction endonucleases may be added to one of or both the ends of the single stranded DNA sequences in the synthesis, or a linker comprising such restriction sites may be ligated to one of or both the ends of the double stranded DNA fragments. Most conveniently such sequences specify multiple recognition

sites for restriction endonucleases (multiple cloning sites MCS).

5 The selected organisms for which artificial promoter libraries can be prepared by the method according to the invention and the various degenerated sequences to be chosen for the promoter libraries of specific organisms are the same as discussed above for the artificial promoter libraries per se.

10

With respect to possible uses of the artificial promoter libraries described above, the invention further provides a method of optimizing the expression of a gene in a microorganism, which comprises

15

a) selecting a set of promoters covering a range of promoter activities in relatively small steps of activity change from an artificial promoter library according to any one of claims 1-26 or constructed by the method according to any one of claims 30-55;

20

b) cloning said set of promoters into said organism placing in each clone said gene under the control of at least one promoter from the set;

25

c) growing the selected clones and screening them to find the one showing optimized flux of product formation.

30 This method is preferably used with organisms selected from the group consisting of prokaryotic and eukaryotic microorganisms such as bacteria, yeasts, other fungi and mammalian cell lines.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. A library of artificial promoters for *L. lactis*, from example 1. The promoter activities (Miller units) were assayed from the expression of a reporter gene (*lacLM*) encoding  $\beta$ -galactosidase transcribed from the different synthetic promoter clones, on the promoter cloning vector pAK80, in strain MG1363, grown in GM17 medium supplemented with 2  $\mu$ g/ml erythromycin. The patterns of the data points indicate which promoter clones differed in either the -35 or the -10 consensus sequence, or in the length of the spacer between these two consensus sequences.

Figure 2. A library of artificial heatshock regulated promoters for *L. lactis*, from example 2. The promoter activities (Miller units) were assayed from the expression of a reporter gene (*gusA*) encoding  $\beta$ -glucuronidase, transcribed from the different synthetic promoter clones, on the chromosome in strain LB436. The cells were grown in GM17 medium supplemented with 2  $\mu$ g/ml erythromycin, and at two different temperatures, 30 and 37 °C. The assay for  $\beta$ -glucuronidase was parallel to the  $\beta$ -galactosidase assay (see example 1) except that X-gluc was used as the substrate.

Figure 3. The library of artificial promoters from example 1, assayed for activity in *E. coli*. The promoter activities (Miller units) were assayed from the expression of a reporter gene (*lacLM*) encoding  $\beta$ -galactosidase transcribed from the different synthetic promoter clones, on the promoter cloning vector pAK80, in strain BOE270, grown in LB medium supplemented with 200  $\mu$ g/ml erythromycin. The patterns of the data points indicate which promoter clones differed in either the -35 or the -10 con-

sensus sequence, or in the length of the spacer between these two consensus sequences.

Figure 4. A library of artificial promoters for *S. cerevisiae*, from example 7. The promoter activities (Miller units) were assayed from the expression of a reporter gene (*lacZ*) encoding  $\beta$ -galactosidase transcribed from the different synthetic promoter clones, on the promoter cloning vector pYLZ-2 in *S. cerevisiae* strain SG24 (URA3-52), grown in SD minimal medium supplemented with 2% glucose. YP24 and YP435 has a 1 bp deletion and a point mutation, respectively, in the UASGCN4p binding site. pTK101 contains a promoter in which the UAS sequence has been deleted and the TATA box is present.

15

Figure 5. Regulation of artificial yeast promoters by external arginine, from example 7. The promoter activities (Miller units) were assayed from the expression of a reporter gene (*lacZ*) encoding  $\beta$ -galactosidase transcribed from the different synthetic promoter clones, on the promoter cloning vector pYLZ-2 in *S. cerevisiae* strain SG24 (URA3-52), grown in SD minimal medium supplemented with 2% glucose, with (SDG) or without 1 g/l arginine (SDG+arg). YP183 has a 13 bp deletion in the binding site for the argR repressor.

25

Figure 6. Regulation of artificial yeast promoters by external amino acids, from example 7. The promoter activities (Miller units) were assayed from the expression of a reporter gene (*lacZ*) encoding  $\beta$ -galactosidase transcribed from the different synthetic promoter clones, on the promoter cloning vector pYLZ-2 in *S. cerevisiae* strain SG24 (URA3-52), grown in SD minimal medium supplemented with 2% glucose (SDG) or complex medium (containing amino ac-

30

ids) without uracil (SC-ura). YP24 has a 1 bp deletion in the UASGCN4p binding site.

#### DETAILED DESCRIPTION OF THE INVENTION

5

In our method, degenerated oligonucleotides are synthesized for the organism or group of organisms for which a promoter library should be constructed. The sequences of the oligonucleotides are written by combining the available knowledge from the literature, on the features that makes a promoter function efficiently in that particular organism. The amount of information that needs to be fixed in the oligonucleotide is somewhat variable among different organisms. In *E. coli* for instance, promoters of considerable strength may be formed by less perfect matches in the -35 and -10 consensus sequences and by spacing between these sequences deviating from 17bp (TTGACA and TATAAT respectively), whereas the requirements for strong promoters in *L. lactis* appear to be more strict.

20

Secondly, the single stranded oligonucleotides are converted into double stranded DNA fragments and cloned into a promoter probing vector. The promoter-containing clones are identified e.g. by their ability to give colonies with various extents of colour on indicator plates. This should in principle give us only very strong promoters, but we discovered that by allowing the spacer sequences between the consensus sequences to vary in a random manner, the strength of the resulting promoters are modulated. In fact, using this method we obtained promoter libraries, spanning the entire range of promoter activities that is likely to become of interest, in small steps of activity change.

35

Optimization of gene expression could then be achieved as follows: 1) From the promoter library one chooses promoters that have e.g. 25%, 50%, 100%, 200% and 400% of the strength of the wild-type promoter. 2) Then, these promoters are cloned in place of the wild-type promoter upstream of the gene of interest. 3) The magnitude of the variable to be optimized (e.g. the flux through a pathway) that is obtained with each of these five constructs is then measured and the optimal construct is used directly as the production organism. It may be necessary to fine-tune the expression further or to expand the range of promoter activities. A direct advantage of this system over the inducible systems described above is that once the optimal promoter activity has been determined, the strain is in principle ready for use in the fermentation process.

In one preferred embodiment the random spacers method of the invention is used for generating a series of constitutive promoters for the Gram-positive bacterium, *Lactococcus lactis*. In other preferred embodiments we show that promoters generated by the random spacers method of the invention are functional in at least two species of Gram-negative bacteria, *Escherichia coli* and *Pseudomonas*, as well as in the Gram-positive bacterium *Bacillus subtilis*. We also show that the strength of the individual promoter is dependent on which organism it is being used in, i.e. that in some organisms a particular promoter is strong, in others it is weak, but in all the organisms tested, the promoters cover a broad range of activities, in small steps of activity change.

Often it is desirable to activate gene expression to a certain extent and at a certain stage of a fermentation, e.g. because the gene product that is expressed inhibits the growth of the cells. It is then useful to combine the



above technique for obtaining promoters of different strength with some regulatory mechanism, e.g. so that the promoter will be activated by a change in the pH, temperature or growth phase.

5

Thus, in another preferred embodiment the random spacers method of the invention is used for generating a series of specifically regulated promoters. As illustrated in Example 2, the above approach is used in combination with  
10 specific regulatory DNA sequences to generate a library of heatshock-regulated promoters for the Gram-positive bacterium, *Lactococcus lactis*.

In addition to prokaryotes, eukaryotic microorganisms  
15 (yeast and other fungi as well as mammalian cell lines,) are important microorganisms for production of a range of organic compounds and various proteins. It is therefore of interest to develop the above approach for modulating gene expression in these organisms as well. Thus, in yet  
20 another preferred embodiment, as illustrated in Example 7, the random spacers method of the invention is used for generating a series of promoters for the bakers yeast, *Saccharomyces cerevisiae*. The promoters are here equipped with GCN4p and ARGR regulation.

25

The regulation of transcription initiation in the eukaryotic cell is somewhat more complex compared to the prokaryote. The transcription start site is normally preceded by a so-called TATA box that contains the consensus  
30 sequence TATAAA or parts hereof, but unlike in the prokaryote, the distance from the TATA box to the transcription start site is much less defined. In *Saccharomyces cerevisiae* this distance is typically 40-120 nucleotides (Oliver and Warmington, 1989). The so-called -35 consensus hexamer which is found in many prokaryotic promoters  
35 is absent in *Saccharomyces cerevisiae*. Instead so-called

upstream activation sequences (UAS) are found upstream of the transcription initiation site. These UAS are recognised by specific DNA binding proteins that can then act as activators of transcription initiation. For instance, the UAS sequence that is found upstream of the genes involved in aminoacid biosynthesis, UASGCN4p, consists of a DNA sequence that specifies a binding site for the GCN4 protein, which activates the transcription of these genes (Hinnebusch, 1992). In contrast to prokaryotes, where the distance between the promoter elements appears to be critical for the strength of promoters (see example 1), the distance between the TATA box and the UAS sequence in eukaryotic promoters is highly variable and may be up to about 1000 bp. Some genes even contain more than one copy of the UAS, but one seems to be sufficient for activation.

One of the UAS sequences known from yeast is the binding site for the GCN4 protein. Promoters that contain the binding site for GCN4 protein should be regulated by the status of amino acid supply in the growth medium: in the absence of amino acids, the GCN4 protein is formed and binds specific UAS sequences to stimulate transcription at promoters involved in biosynthesis of aminoacids. The consensus sequence for the GCN4 protein binding site (UASGCN4p) is a short inverted repeat, TGACTCA.

A promoter in *Saccharomyces cerevisiae* that is activated by the GCN4 protein is the ARG8 promoter. In this promoter, there is only one copy of the UASGCN4p sequence, and it is located 59 bp from the TATA box (we refer to this distance as spacer 1). Transcription initiation takes place some 40-60 nucleotides downstream of the TATA box. The ARG8 promoter also contains a DNA sequence that functions as the binding site for the arginine repressor, *argR* (Crabeel et al., 1995), which makes the promoter four fold repressible by external arginine.

Thus, in this case the promoter is located within 136 bp, and it contains two regulatory features, which makes the system attractive for developing a promoter library by the random spacers method of the invention as outlined in the previous examples. But the method is not limited to this model system; in principle, any combinations of TATA boxes, UAS sequences, repressor binding sites etc., separated by spacers smaller than about 1000 nucleotides, should be suitable as a starting point for this method.

#### EXAMPLE 1

Design of a degenerated oligonucleotide for a *L. lactis* promoter library.

According to the literature (see review in de Vos & Simons, 1994), strong promoters in *L. lactis* tend to have the following nucleotide sequences in common (numbers refer to the position relative to the transcription initiation site, which is given number +1): -12 to -7: TATAAT; -15 to -14: TG; -35 to -30: TTGACA. The spacing between -10 and -35 seems to be 17 nucleotides. However, closer comparison of the promoter sequences that have been published for *L. lactis* reveals that in a number of positions besides the ones mentioned above, nucleotides are more or less well conserved. Some of these positions are: -1: A; -3: A or T (=W); -6: A; -13: A or G (=R); -40 to -36: TATTC. In addition, Nilsson and Johansen (1994, BBA) pointed out two motives, +1 to +8: GTACTGTT, and -44 to -41: AGTT, that appear to be well conserved between relatively strong promoters (promoters for transfer RNA and ribosomal RNA operons) from *L. lactis*. These motives may confer both strength and growth rate dependent expression from the promoter.

When these additional motives are included, one arrives at the following 53 nucleobase degenerated sequence for

an efficient promoter in *L. lactis*. Out of these 53 nucleobases, 34 bases are conserved, two are semi-conserved (R and W) and 17 are allowed to vary randomly between the four nucleobases.

5

5' AGTTTATTCTTGACANNNNNNNNNNNNNNTGRTATAATANNWNAGTACTGTT 3'

In addition, this degenerated sequence is flanked by sequences that specify multiple recognition sites for restriction endonucleases (multiple cloning site MCS), and the actual oligonucleotide mixture to be synthesized has the following degenerated sequence reported in SEQ ID No. 1:

15

MboI  
DpnI  
AlwI  
NlaIV  
BstYI  
BamHI MseI  
AlwI AflIII SspI NsiI

20

1  
5'CGGGATCCTTAAGAATATTATGCATNNNNNAGTTTATTCTTGACANNNNNNNNNNNNNNT

25

AluI  
PvuII  
NspBII  
SfiI  
MseI Fnu4HI  
RsaI HpaI PstI  
ScaI HincII BbvI EcoRI

30

61 100  
GRTATAATANNWNAGTACTGTTAACTGCAGCTGAATTCGG 3'

35

A mixture of oligonucleotides according to this specification was synthesized by Hobolth DNA synthesis.

This oligonucleotide mixture is single stranded initially and must be converted into double stranded DNA fragments for the purpose of cloning. This was done by synthesizing in vitro a 10 bp oligonucleotide, having a sequence complementary to the 3' end of the promoter oligonucleotide. This oligonucleotide was then used as a primer for second strand synthesis by the Klenow fragment of DNA polymerase

45

I in the presence of dATP, dCTP, dGTP and dTTP. In this way the second DNA strand became exactly complementary to the first DNA strand.

5 The result is a mixture of 100 bp double stranded DNA fragments containing multiple recognition sites for restriction endonucleases in both the 3' and 5' end. These DNA fragments were then cut with restriction endonucleases in order to create suitable "sticky" ends compatible  
10 with the ends of the vector DNA fragment, pAK80 (Israelsen *et al.*, 1995). pAK80 is a shuttle vector, meaning that it has replication origins for propagation in both *E. coli* and *L. lactis*. In this way, the cloning steps can be conveniently performed in *E. coli*, while the  
15 subsequent physiological experiments can be done in *L. lactis*. Furthermore, pAK80 carries a promoterless  $\beta$ -galactosidase reporter gene system (*lacLM*) downstream a multiple cloning site. Thus, pAK80 does not express the *lacLM* genes, unless a promoter is inserted in the multi-  
20 ple cloning site.

Two cloning strategies were used for cloning the mixture of double stranded DNA fragments into the cloning vector pAK80:

25

1) The mixture was digested with *BamHI* and *PstI* and the vector pAK80 with *PstI* and *BglII* (*BglII* is compatible with *BamHI*).

30

2) The mixture was digested with *SspI* and *HincII* and the vector pAK80 with *SmaI* (all three enzymes produce blunt end DNA fragments).

35 In both cases the vector DNA was further treated with Calf Intestine Phosphatase (CIP) to prevent religation of the cloning vector. Subsequently, the fragment mixture

and vector DNA were ligated overnight at 16 °C using T4 DNA ligase and standard ligation conditions.

Ligation mixtures were transformed into *E. coli* K-12  $\Delta$ lac, with selection for erythromycin resistance. Cells of the *E. coli* K-12 strain, MT102, were made competent using standard treatment with  $\text{Ca}^{++}$  ions (Maniatis et al., 1982). Ligation mixtures were then transformed into these cells using a standard transformation procedure (Maniatis et al., 1982), and the resulting clones were screened for  $\beta$ -galactosidase activity that will produce blue colonies on plates containing X-gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside). The transformation mixture was plated on LB plates containing 200  $\mu\text{g}/\text{ml}$  erythromycin, 1% glycerol and 100  $\mu\text{g}/\text{ml}$  X-gal. 150 erythromycin resistant transformants were obtained, all white initially, but after prolonged incubation (two weeks at 4 °C), 46 of these colonies had become light blue. Thus, using strategy 1) we found 17 blue colonies (CP30 to CP46), and using strategy 2) we found 29 blue colonies (CP1 to CP29).

Plasmid DNA was isolated from each of these clones (CP1 to CP46) and analysed by restriction enzyme mapping. Nearly all plasmids (except CP31 and CP43) contained promoter fragments inserted into the MCS of pAK80, in the orientation that would direct transcription of the otherwise promoterless *lacLM* genes on this vector.

These 46 plasmid DNA preparations were then transformed into *L. lactis* subspecies *lactis* MG1363 with selection for erythromycin resistance. Cells of the *L. lactis* subspecies *lactis* strain, MG1363 (Gasson, 1983) were made competent by growing the cells overnight in SGM17 medium, containing 2% glycine, as described by Holo and Nes (1989). Plasmid DNA from each of the 46 clones described above was then transformed into these cells using the electroporation procedure (Holo and Nes, 1989). After re-

generation, the cells were plated on SR plates containing 2 µg/ml erythromycin. Subsequent screening for blue color on X-gal plates revealed large differences in β-galactosidase activity between the 46 clones; some clones gave dark blue colonies after 24 hours of incubation, others only light blue colonies after more than 1 week of incubation.

The β-galactosidase activities of liquid cultures of the 46 clones in MG1363 were then determined as described by Miller (1972) and modified by Israelsen et al. (1995). Cultures of the strain MG1363, each carrying one of the 46 plasmid derivatives of pAK80, were grown in M17 medium supplemented with 1% glucose overnight at 30 °C. 25-100 µl of these cultures were then used in the subsequent β-galactosidase assay, except in the case of the weakest promoter clones, where 2 ml of culture was used (after 20 fold concentration by centrifugation). These results are shown in Figure 1. Apparently, there are very large differences in the efficiency of the cloned promoter fragments, and together these clones cover a range of promoter activities from 0.3 units of β-galactosidase activity to more than 2000 units, which is probably the strongest promoters known for this organism.

In addition, the broad range is covered by small changes in activity and, therefore, these promoter fragments will allow us not only to obtain a wide range of expression of genes in *L. lactis*, but also to tune the expression of genes in *L. lactis* in small steps for the purpose of flux optimization.

DNA sequencing of the 46 clones described above revealed that most of the inserted promoter fragments had the DNA sequence that was originally specified for the oligonucleotide design (see above), whereas the sequence of the remaining fragments deviated slightly from that sequence.

Most of the promoter fragments that gave the lower activities in the  $\beta$ -galactosidase assay (70 units of  $\beta$ -galactosidase or less) had either an error in one of the consensus sequences or a 16 bp spacer between the consensus sequences. This result is in accordance with the dogma, i.e. that changes in the consensus sequences have strong effects on the activity of a given promoter, and emphasizes the fact that a more subtle approach is needed in order to generate a promoter library that covers a range of activities in small steps of activity change. Clearly, if we would have allowed only changes in the consensus sequences and/or changes in the length of the spacer, instead of allowing the sequence of the spacers to vary randomly, only fairly weak promoter clones would have resulted, and the resulting library would not be suitable for metabolic engineering.

In general, the clones that gave high activities (more than 70 units) had the same sequence as specified by the oligonucleotide. In total, the activity of the clones that had intact consensus sequences and 17 basepair spacer length, spanned activities from 5 units (CP4) to 2000 units (CP25). These results therefore show that at least a 400 fold change in promoter activity can be obtained by randomizing the spacer while the consensus sequences are kept constant.

Usually, for metabolic engineering purposes, relatively strong promoters are desired. However, there may also be cases where rather weak promoters are needed. The relatively few errors that had occurred during synthesis of the above oligonucleotide mixture, were not intended to be present in the promoter fragments; and our data then suggest that it may be useful to generate, deliberately, a mixture of oligonucleotides which have a low percentage of errors in the consensus sequences.



Enzymes used in the various enzymatic reactions above were obtained from and used as recommended by Pharmacia and Boehringer.

5   **EXAMPLE 2**

**Design of a degenerated oligonucleotide for a library of temperature regulated *L. lactis* promoters.**

- 10   This example illustrates the development of a temperature regulated promoter library for *L. lactis*. A regulatory element comprising an eight basepair inverted repeat that has been shown to be involved in the heatshock response of *L. lactis* is inserted a few base pairs upstream of the  
15   -35 sequence. The minimal extent of such a regulatory element seems to be 27 basepairs:

5'-TTAGCACTCNNNNNNNNNGAGTGCTAA-3'

IR       spacer       IR

20

- containing a 9 bp (or longer) inverted repeat (IR) separated by 9 (or fewer) basepairs. It should therefore be possible to combine this inverted repeat with the approach for obtaining constitutive promoters of different  
25   strength and thus obtain a series of promoters with various basal activities which can be induced several fold by changing the temperature of the culture medium.

- Therefore, an oligonucleotide was designed, which includes the core part (from position -35 to +6) of the sequence from the constitutive promoter series above (see  
30   example 1 and SEQ ID No. 1). The sequence upstream of the -35 hexamer has been replaced by the above inverted repeat sequence, and the sequence downstream of position +6  
35   has also been modified, eliminating two conserved regions compared to example 1 (+1 to +8: GTACTGTT and -44 to -41: AGTT, which have been implicated in growth rate regula-

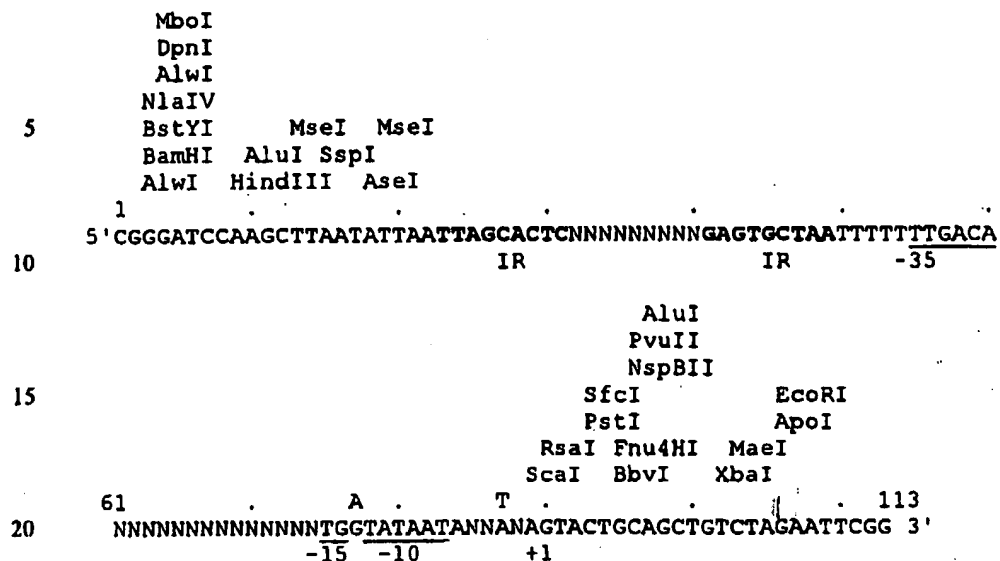
tion, but which turned out to be dispensable with respect to creating strong promoters for *L. lactis*, see example 1). The sequence of the spacer (sp.1) between the two inverted DNA sequences in the inverted repeat was here allowed to vary randomly in order to see whether this had any effect on the temperature regulation of the resulting promoters, e.g. how many fold they could be induced by changing the temperature. The importance of the spacing (sp.2) between the inverted repeat and the -35 hexamer is not known, but in principle this region may contribute to or modulate the heatshock response of promoters. In order to limit the number of parameters, however, we have chosen here to include a naturally occurring configuration (derived from the *dnaJ* promoter from *L. lactis*; van As-seldonk et al., 1993): a short spacer sequence consisting of 5 times T.

When these sequences are combined, one arrives at the following 73 bp consensus sequence for a temperature regulated promoter in *L. lactis*. Out of these 73 bp, 45 are conserved, two are semi-conserved (R and W) and 26 are allowed to vary randomly between the four nucleobases.

5' TTAGCACTCNNNNNNNNNGAGTGCTAATTTTTTTGACNNNNNNNNNNNNNTGR  
IR spacer 1 IR sp.2 spacer 3

TATAATANNWNAGTACTG 3'

In addition, this degenerated sequence was flanked by sequences that specify multiple recognition sites for restriction endonucleases (multiple cloning sites MCS), and the actual oligonucleotide mixture that is being synthesized has the following degenerated sequence reported in SEQ ID No. 2:



This oligonucleotide mixture was converted into double stranded DNA fragments (DSDF) and cloned into the promoter cloning vector pLB85i. pLB85i has an origin for replication in *E. coli* but not in *L. lactis* and selectable markers for both organisms. Instead of an origin of replication for *L. lactis*, it contains the *attP* sequence which can direct the insertion of the plasmid into the *L. lactis* chromosome if the *int* gene product is supplied *in trans*. This example therefore also serves to illustrate the modulation of chromosomally encoded genes, using promoters generated through the random spacers method. pLB85i is also a promoter cloning vector and contains a multible cloning site upstream of a promoterless *gusA* gene. *gusA* is similar to the *lacZ* and *lacLM* screening systems used in the other examples, except that the substrate is X-gluc instead of X-gal. It was chosen as the reporter gene for this particular application, where heatshock regulated promoters are supposed to be analysed. This was because the *gusA* gene product is not heat labile, which seems to be a problem connected with the *lacLM* gene used in example 1.

The DSDF mixture was here digested with *XbaI* and *AseI* and the vector pLB85i was digested with *XbaI* and *NdeI* (*NdeI* is compatible with *AseI*) and further treated with alkaline phosphatase to remove the 5 phosphate groups from the vector. After ligation of the DFDS mixture and the vector, the ligation mixture was transformed into strain KW1, a *gusA* negative *E. coli* strain, with selection for ampicillin resistance. This resulted in approximately 100 colonies of which 80% had different blue color intensities, indicating that these clones were carrying putative heatshock promoter fragments. 20 blue clones were picked for further analysis. Restriction analysis of plasmid preparations from these clones showed that they had inserts of approximately the right size. Subsequently, 6 of these plasmid preparations were used to transform LB436 (a derivative of *L. lactis* MG1363 which contain a second plasmid, pLB81, supplying the necessary *int* gene for integration of the plasmids into the *attB* site on the *L. lactis* chromosome). From each of the transformations colonies were isolated with putative integrations of the constructs into the *attB* attachment site on the *L. lactis* chromosome. The integration of the constructs was confirmed by standard colony PCR, using one primer in the *attB* region and one primer in the pLB85i region. The clones were subsequently tested for their ability to form blue colonies on GM17 medium supplemented with erythromycin and 100 µg/ml X-gluc at 30 °C and 37 °C, respectively, and the clones showed a clear difference in color intensity at the two temperatures, indicating that the promoter activities were now under temperature regulation. Next the *gusA* expression from the heatshock promoters was measured in liquid cultures at the two temperatures, for 5 selected heatshock-regulated promoters, see Figure 2. The clones had different activities and covered a broad range of promoter activities at 30 °C,

and 4 out of the 5 clones (all except HP6) gave higher promoter activities at 37 °C, which shows that the promoters were indeed temperature regulated. Interestingly, the promoters were regulated by almost the same fold, i.e. from 1.7 to 2.3 fold by the temperature shift from 30 to 37 °C, which indicates that the spacer (spacer 1) in the inverted repeat is of minor importance with respect to determining the fold of induction of these artificial promoters.

10

We have here been looking at the accumulated *gusA* activity, and although the data clearly show that the promoters are temperature regulated, one skilled in the art of analysing gene expression will appreciate that this makes it more difficult to observe the changes in promoter activities, that are brought about for instance by some change in external parameter such as temperature. In addition, it has been shown that the activity of heatshock regulated promoters is temporarily ten fold higher immediately after the temperature shift, than it appears when the steady state levels are compared as they are on Figure 2. To observe the change in promoter activity more carefully, one should look at the rate of protein or mRNA synthesis, before and at various times after the perturbation in temperature. We therefore isolated RNA from the 5 clones at various times after the change in temperature from 30 °C to 37 °C and visualized the *gusA* mRNA by standard Northern blotting (Maniatis et al., 1983), to analyze how many fold the individual promoter clone was induced by the temperature change.

30

### EXAMPLE 3

The Gram-positive bacterium, *Bacillus subtilis* is used extensively as an industrial bioreactor for the produc-

35

tion of a range of heterologous proteins. It was therefore of interest to test whether the random spacers method of the invention could also be used to generate promoter libraries for this organism. The consensus sequences for *Bacillus subtilis* are very similar to the consensus sequences for *E. coli* and *L. lactis*, and we could therefore test whether the approach was also valid for *Bacillus subtilis* by subcloning a number of CP promoters into a promoter cloning vector for this *Bacillus subtilis* and then ask 1) whether the CP promoters are active in *Bacillus subtilis* and 2) whether, also in this organism, the spacer between the consensus sequences plays an important role for the promoter strength. We chose to use the promoter cloning vector, pDG268, which is designed for integration of promoter fusions to *lacZ* into the *amy* locus on the *Bacillus subtilis* chromosome. The vector confers ampicillin and neomycin resistance, and it will replicate in *E. coli* for the initial cloning purposes, but not in *Bacillus subtilis*. When the linearized form of the vector is transformed into *Bacillus subtilis* it will be inserted into the *amy* locus on the *Bacillus subtilis* chromosome. This example therefore also serves as an example of the use of the promoters generated through the random spacers methods for modulation of chromosomally encoded genes.

The CP promoters were indeed active also in *Bacillus subtilis*, and the individual strength of the promoters were again very different from each other. The fact that the promoter library also covered a broad range of promoter activities in *Bacillus subtilis*, shows that the random spacers method is also valid for this organism.

**EXAMPLE 4**

The CP promoters of example 1 were designed for use in *L. lactis*, but the consensus sequences of *E. coli* promoters are included in the sequence of the oligonucleotide given in example 1. Furthermore, the vector used for generating the CP promoter library in example 1 is a shuttle vector for *L. lactis* and *E. coli*, so this also allowed us to analyse the activity of the CP promoters in the Gram-negative bacterial host, *E. coli*. Figure 3 shows the activity of 33 of the CP promoters in *E. coli*. Clearly, the activity of the individual CP promoters is also here very different, and together the promoters cover a broad range of activities. Interestingly, the correlation between the strength of the individual promoters in *E. coli* and *L. lactis* was not very strong: some promoters that were found to be strong promoters in *L. lactis* were found to be weak in *E. coli* and vice versa.

The activity of the promoters, in terms of  $\beta$ -galactosidase activity, was generally much lower than the activity found in *L. lactis*. This was probably due to the fact that the promoter cloning vector, pAK80 is designed and optimised for use in the Gram-positive bacterium, *L. lactis*, and the translation efficiency in *E. coli* could therefore be low. We therefore sub-cloned three of the CP promoters (CP20, CP22 and CP25) into the promoter cloning vector, pCB267, which was designed for cloning *E. coli* promoters in upstream of a promoterless *lacZ* gene, encoding  $\beta$ -galactosidase. In this promoter system, the CP promoters turned out to be very efficient promoters, but the relative difference in strength between the three promoters was conserved. Thus, the CP25 promoter gave 2.5 fold higher activity than the hybrid promoter, ptac, which is considered to be among the strongest promoters known for

use in *E. coli*. These data therefore further demonstrate the strength of our approach: by analysing a relatively small number of promoter clones obtained through the random spacers method of the invention, we have managed to  
5 arrive at some of the strongest promoters in both *E. coli* and *L. lactis*.

#### EXAMPLE 5

10 Bacteria belonging to the Gram-negative species, *Pseudomonas*, are becoming increasingly important due to their application in e.g. biodegradation of chemical waste products in polluted soil. But also here the genetic engineering is hampered by the lack of suitable promoters  
15 and expression systems. The literature on *Pseudomonas* promoters revealed that the consensus sequences for *Pseudomonas* are somewhat less well defined than those of *E. coli*, *L. lactis* and *Bacillus subtilis*. Thus, in the -35 region one often finds TTGR conserved (R=A or G) whereas  
20 the rest of the -35 consensus sequence is varying between different promoters. The -10 consensus sequence is probably TATRAT. The spacing between the TTGR motive and the -10 sequence is 18-19 bp, which is equivalent of the 16-17 bp spacer often found in *E. coli*. It follows that the  
25 consensus sequences for vegetative promoters in this organism are quite close to the consensus sequences for *E. coli* and *L. lactis*.

We therefore tested a range of the CP promoters from example 1 in *Pseudomonas putida* by cloning the promoters  
30 into a cloning vector that contains a promoterless  $\beta$ -galactosidase gene. Again the activity of the CP promoters differed in strength over a broad range of promoter activities. These results show that also in this species,  
35 the random spacers method of the invention could be used



to generate both relatively strong promoters and also a broad range of promoter activities covered in small steps of activity change

#### 5    **EXAMPLE 6**

We also designed an oligonucleotide based on the consensus sequences stated in example 5 and incorporated multiple cloning sites as described in example 1. The following oligonucleotide was synthesised.

MCS-(N)<sub>8</sub>-TTGR-N<sub>19</sub>-TATRAT-(N)<sub>8</sub>-MCS

The oligonucleotide was converted to double stranded DNA, using a primer homologous to the 3' end of the oligonucleotide and cloned upstream of a promoterless  $\beta$ -galactosidase gene on a promoter probe vector. The ligation mixture was transformed directly into a *Pseudomonas putida* strain with selection for the antibiotics resistance carried by the plasmid and on plates containing X-gal. This resulted in approximately 100 clones of various blue color intensities. Subsequently, 30 clones were analysed for  $\beta$ -galactosidase activity as described above. These results showed that also in this species, the random spacers method of the invention could be used to generate both relatively strong promoters and also a broad range of promoter activities covered in small steps of activity change.

#### 30    **EXAMPLE 7**

Design of a degenerated oligonucleotide for a *Saccharomyces cerevisiae* promoter library.

35    A 199 bp oligonucleotide was designed, which includes, starting from the 5' end: an *EcoRI* restriction site (for

```

      PleI
EcoRI   HinfI
20     ApoI   MaeIII
       1 . . . . .
5' CAGAATTCGTGAGCTCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
    UAScch4p
76
25     61 . . . . .
    NNNNNNNNNNNNNNNNTATAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
        TATA-box
          MseI                               TfiI
30     AflII           MaeIII             HinfI
       121 . . . . .
NCTCTTAAGTGCAAGTGACTGCGAACATTTTTTTCGTTGTTAGAATAATTCAAGAATCG
         TI box
35                MboI
                   DpnI
                   AlwI
                   NlaIV
                   BstYI
40                BamHI
                   AlwI
                   NlaIII
       181               199
CTACCAATCATGGATCCCG 3'
45                ARGStart

```

09/14/2004, EAST Version: 1.4.1

complementary to the last 23 bp of the 3' end of the 199 bp degenerated oligonucleotide as described in example 1. Subsequently, it was cloned into either of the two promoter cloning vectors, pYLZ-2 and pYLZ-6 (Hermann et al., 1992), as follows: the DSDF mixture and the vector were both double-digested with *EcoRI* and *BamHI*, and the DSDF were ligated to the vector DNA. The ligation mixture was transformed into *E. coli* as described in example 1, with selection for ampicillin resistance. Plasmid DNA was isolated from 500 individual clones and screened for the presence of putative promoter fragments by digestion with the restriction enzymes, *EcoRI* and *BamHI*.

17 clones were found to have an *EcoRI* - *BamHI* insert of approximately 200 bp. Plasmid DNA from these 17 clones were transformed into *S. cerevisiae*, with selection for the URA3 marker carried by the plasmids and assayed for production of  $\beta$ -galactosidase. Figure 4 shows the resulting activities of  $\beta$ -galactosidase for the 17 clones. All the promoters have higher activities than the cloning vector itself without promoter fragment inserted (pYLZ-2). More important however, also in this case the clones cover a range of promoter activities in small steps of activity change.

Sequence analysis revealed that the 17 clones discussed above had a perfect TATA box (TATAAA) between the spacer 1 and spacer 2, while two of the 17 clones, YP24 and YP435 each had a defect in the UASGCN4p. However, the activity of YP435 was 39 units which is close to the value obtained with YP212. These data then suggest that the impact of the random sequence of the spacers on the promoter strength, is stronger than the impact of the state of the UASGCN4p binding site.

As discussed above, the artificial yeast promoters had built in two different regulatory features. One is that the promoters should be regulated by the presence of arginine in the growth medium. To test whether the artificial yeast promoters were also regulated by arginine, we grew a number of clones in minimal medium, with and without arginine (SD + 2% glucose; SD + 2% glucose + 1 g/l arginine). Figure 5 shows the result of these experiments. The clones, YP18, YP212, YP435 were indeed regulated 5, 8, 15 fold, respectively, by the presence of arginine. YP183 was not regulated by arginine, and sequence analysis confirmed that this promoter clone had a 13 bp deletion in the arginine repressor binding site.

We also tested whether the promoters were regulated by external amino acids in the growth medium, by analyzing the promoter activity of some of the yeast promoter clones in minimal medium (SD + 2% glucose) and rich medium (SC + 2% glucose-URA). Figure 6 shows the result of these experiments. Indeed, the promoters present in clones YP18, YP212, YP435, and YP183 were regulated from 2 to 10 fold, by the presence of amino acids. YP24 was not regulated, in accordance with the error that had occurred in the UASGCN4p site on this clone, see above.

The results on amino acid and specific arginine regulation demonstrate that the random spacers method of the invention can be used for generating promoters which cover a broad range of promoter activities and which can be regulated by external signals. The regulatory aspect of the invention is here exemplified by the amino acid and arginine regulation, but is not limited to these cases.

## REFERENCES

- Crabeel, M., de Rijcke, M., Seneca, S., Heimberg, H., Pfeiffer, I., and Matisova, A., 1995. Further definition  
5 of the sequence and position requirements of the arginine control element that mediates repression and induction by arginine in *Saccharomyces cerevisiae*. *Yeast* 11: 1367-1380
- de Vos, W.M., and Simons, G., 1994. Gene cloning and expression systems in lactococci, p.52-105. In M.J. Gasson and W.M. de Vos (eds.), *Genetics and biotechnology of lactic acid bacteria*. Blackie Academic & Professional, Glasgow, United Kingdom.
- 10
- Gasson, M.J. 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. *J. Bacteriol.*, 154, 1-9.
- 15
- Guarente, L., 1983. Yeast promoters and *lacZ* fusions designed to study expression of cloned genes in yeast. *Methods Enzymol.* 101, 181-191.
- 20
- Hermann, H., Hacker, U., Bandlow, W., and Magdolen, V., 1992. pYL2 vectors: *Saccharomyces cerevisiae*/*Escherichia coli* shuttle plasmids to analyze yeast promoters. *Gene* 119: 137-41.
- 25
- Hinnebusch, A.G., 1992. General and Pathway-specific Regulatory Mechanisms Controlling the Synthesis of Amino Acid Biosynthetic Enzymes in *Saccharomyces cerevisiae*, p. 319-414. In E.W. Jones, J.R. Pringle and J.R. Broach (eds.), *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- 30
- 35
- Holo, H., and Nes, I.F., 1989. High frequency transformation, by electroporation, of *Lactococcus lactis* subsp.

*cremoris* grown with glycine in osmotically stabilized media. *Appl. Environ. Microbiol.*, 55, 3119-3123.

5 Israelsen, H. 1995. Cloning and partial characterization of regulated promoters from *Lactococcus lactis* Tn917-lacZ integrants with the new promoter probe vector, pAK80. *Appl. Environ. Microbiol.*, 61, 2540-2547.

10 Jensen, P.R., Westerhoff, H.V., and Michelsen, O., 1993. Excess capacity of  $H^+$ -ATPase and inverse respiratory control in *Escherichia coli*. *EMBO J.*, 12, 1277-1282.

Kacser, H. and Burns, J.A. 1973. The control of flux. *Symp. Soc. Exp. Biol.*, 27: 65-104.

15 Maniatis, T., Fritsch, E.F., and Sambrook, J., 1982. *Molecular cloning*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

20 Miller, J.H., 1972. *Experiments in molecular genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

25 Nilsson, D., and Johansen, E., 1994. A conserved sequence in tRNA and rRNA promoters of *Lactococcus lactis*. *Biochim. Biophys. Acta*, 1219: 141-144.

Oliver, S.G., and Warmington, J.R., 1989. Transcription. In *The yeasts, volume 3*, Rose and Harrison (Eds.). Academic Press, London, 117-160.

30

Schaaff, I., Heinisch, J., Zimmermann, F.K., 1989. Overproduction of glycolytic enzymes in yeast. *Yeast*, 5: 285-290.

35 van Asseldonk, M., Simons, A., Visser, H., de Vos, W.M., and Simons, G., 1993. Cloning, nucleotide sequence, and

regulatory analysis of the *Lactococcus lactis* dnaJ gene.  
*J. Bacteriol.* 175, 1637-1644.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT:
  - (A) NAME: Peter Ruhdal Jensen
  - (B) STREET: Soegaardsvej 19
  - (C) CITY: Gentofte
  - (E) COUNTRY: Denmark
  - (F) POSTAL CODE (ZIP): DK-2820
- (ii) TITLE OF INVENTION: Artificial promoter libraries for selected organisms and promoters derived from such libraries
- (iii) NUMBER OF SEQUENCES: 58
- (iv) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
- (vi) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: DK 886/96
  - (B) FILING DATE: 23-AUG-1996

## (2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 100 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactococcus lactis
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION: 26..82
  - (C) IDENTIFICATION METHOD: experimental
  - (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Artificial promoter library"  
/note= "A degenerated sequence specifying a mixture of artificial promoters covering a wide range of expression in small steps in L. lactis"
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 31..45
  - (D) OTHER INFORMATION: /standard\_name= "Consensus sequence"
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature



- (B) LOCATION: 60..69
- (D) OTHER INFORMATION: /standard\_name= "Consensus sequence"

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 74..82
- (D) OTHER INFORMATION: /standard\_name= "Consensus sequence"

(ix) FEATURE:

- (A) NAME/KEY: -35\_signal
- (B) LOCATION: 40..45
- (D) OTHER INFORMATION: /standard\_name= "-35 box"

(ix) FEATURE:

- (A) NAME/KEY: -10\_signal
- (B) LOCATION: 63..68
- (D) OTHER INFORMATION: /standard\_name= "Pribnow box"

(ix) FEATURE:

- (A) NAME/KEY: misc\_recomb
- (B) LOCATION: 3..25
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Multiple cloning site"  
/label= MCS  
/note= "A sequence specifying recognition sites for the restriction endonucleases: NlaIV, BstYI, BamHI, AlwI, MboI, DpnI, AflII, MseI, SspI, NsiI."

(ix) FEATURE:

- (A) NAME/KEY: misc\_recomb
- (B) LOCATION: 74..98
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Multiple cloning site"  
/label= MCS  
/note= "A sequence specifying recognition sites for the restriction endonucleases: ScaI, RsaI, HpaI, HincII, MseI, SfcI, PstI, Fnu4HI, BbvI, PvuII, NspBII, AluI, EcoRI."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CGGGATCCTT AAGAATATTA TGCATNNNNN AGTTTATTCT TGACANNNNN NNNNNNNNT 60

GGTATAATAN NANAGTACTG TTAAGTGCAG CTGAATTCGG 100

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION:23..95
- (D) OTHER INFORMATION:/standard\_name= "Artificial promoter library"  
/note= "A degenerated sequence specifying a mixture of artificial temperature regulated promoters covering a wide range of expression in small steps in *L. lactis*"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION:23..49
- (D) OTHER INFORMATION:/standard\_name= "Sequence providing temperature regulation to promoters"  
/note= "This sequence comprising two inverted repeats separated by a short spacer provides temperature (heat shock) regulation to promoters in Gram-positive bacteria"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION:50..60
- (D) OTHER INFORMATION:/standard\_name= "Consensus sequence"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION:75..84
- (D) OTHER INFORMATION:/standard\_name= "Consensus sequence"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION:89..95
- (D) OTHER INFORMATION:/standard\_name= "Consensus sequence"

## (ix) FEATURE:

- (A) NAME/KEY: -35\_signal
- (B) LOCATION:55..60
- (D) OTHER INFORMATION:/standard\_name= "-35 box"

## (ix) FEATURE:

- (A) NAME/KEY: -10\_signal
- (B) LOCATION:78..83
- (D) OTHER INFORMATION:/standard\_name= "Pribnow box"

## (ix) FEATURE:

- (A) NAME/KEY: misc\_recomb
- (B) LOCATION:3..22
- (D) OTHER INFORMATION:/standard\_name= "Multiple cloning site"  
/label= MCS  
/note= "A sequence specifying recognition sites for the restriction endonucleases: NlaIV, BstYI, BamHI, AlwI, MboI, DpnI, HindIII, AluI, MseI (2 sites), SspI, AseI."

## (ix) FEATURE:

- (A) NAME/KEY: misc\_recomb
- (B) LOCATION:89..111
- (D) OTHER INFORMATION:/standard\_name= "Multiple cloning site"  
/label= MCS  
/note= "A sequence specifying recognition sites for the"

restriction endonucleases: ScaI, RsaI, SfcI, PstI, Fnu4HI, BbvI, PvuII, NspBII, AluI, XbaI, MaeI, EcoRI, ApoI."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
CGGGATCCAA GCTTAATATT AATTAGCACT CNNNNNNNNN GAGTGCTAAT TTTTGTGACA      60
NNNNNNNNNN NNNNTGGTAT AATANNANAG TACTGCAGCT GTCTAGAATT CGG          113
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 199 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

- (A) NAME/KEY: protein\_bind
- (B) LOCATION:10..16
- (D) OTHER INFORMATION:/function= "Activating promoters in *S. cerevisiae*"  
/bound\_moiety= "GCN4 protein"  
/standard\_name= "Upstream activating sequence"  
/label= UAS\_GC4p  
/note= "A DNA sequence that specifies a binding site for the GCN4 protein, which activates the transcription of genes involved in aminoacid synthesis in *S. cerevisiae*."

(ix) FEATURE:

- (A) NAME/KEY: TATA\_signal
- (B) LOCATION:67..72
- (D) OTHER INFORMATION:/standard\_name= "TATA box"

(ix) FEATURE:

- (A) NAME/KEY: misc\_signal
- (B) LOCATION:122..144
- (D) OTHER INFORMATION:/function= "Transcription initiation"  
/standard\_name= "TI box"

(ix) FEATURE:

- (A) NAME/KEY: protein\_bind
- (B) LOCATION:122..144
- (D) OTHER INFORMATION:/bound\_moiety= "Arginine repressor"  
/standard\_name= "arginine repressor binding site"  
/label= argR

(ix) FEATURE:

- (A) NAME/KEY: misc\_RNA
- (B) LOCATION:145..192
- (D) OTHER INFORMATION:/function= "Spacer"  
/standard\_name= "Part of native sequence for ARG8"

gene incl. first codon"

(ix) FEATURE:

- (A) NAME/KEY: misc\_recomb
- (B) LOCATION:3..8
- (D) OTHER INFORMATION:/standard\_name= "Recognition site  
for restriction endonuclease EcoRI"  
/label= EcoRI\_site

(ix) FEATURE:

- (A) NAME/KEY: misc\_recomb
- (B) LOCATION:192..197
- (D) OTHER INFORMATION:/standard\_name= "Recognition site  
for restriction endonuclease BamHI"  
/label= BamHI\_site

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION:10..192
- (D) OTHER INFORMATION:/standard\_name= "Artificial  
promoter library"  
/note= "A degenerated sequence specifying a mixture of  
artificial promoters covering a wide range of expression  
in small steps in *S. cerevisiae*"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

CAGAATTCGT GACTCANNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	60
NNNNNNNNNN NNNNNNTATA AANNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	120
NCTCTTAAGT GCAAGTGACT GCGAACATTT TTTTCGTTG TTAGAATAAT TCAAGAATCG	180
CTACCAATCA TGGATCCCG	199

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Pseudomonas putida*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION:1..45
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION:/evidence= EXPERIMENTAL  
/standard\_name= "Artificial promoter library"  
/note= "A degenerated sequence specifying a mixture of  
artificial promoters covering a wide range of expression  
in small steps in *Pseudomonas putida*"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

NNNNNNNNTT GRNNNNNNNN NNNNNNNNNN NTATRATNNN NNNNN

45

## (2) INFORMATION FOR SEQ ID NO: 5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

## (ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 4..60
- (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp1

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CATACGGGAG TTTATTCTTG ACAGTTCCAC CTCGGGTGA TATAATATCT CAGTACTGTT

60

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

## (ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 4..60
- (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CATGGCTTAG TTTATTCTTG ACAGGGTAGT ATCACTGTGA TATAATAGGA CAGTACTGTT

60

## (2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 59 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Lactococcus lactis*
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION: 4..59
  - (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp11

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CATAAGTGAG TTTATTCTTG ACCCGGACGC CCCCTTTGA TATAATAAGT AGTACTGTT

59

## (2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Lactococcus lactis*
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION: 4..60
  - (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp12

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CATATACAAG TTTATTCTTG AACTAGTCG GCCAAATGA TATAATACCT GAGTACTGTT

60

## (2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional  
promoter"  
/label= Cp13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CATGCTTTAC TTTATTCTTG ACAAACCAC CAGCTTTTGG TATAATACGT GAGAACTGTT 60

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional  
promoter"  
/label= Cp14

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CATGACGGAG TTTATTCTTG ACACAGGTAT GGAATTATGA TATAATAAAA CAGTACTGTT 60

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION:4..60

(D) OTHER INFORMATION:/standard\_name= "Constitutional  
promoter"  
/label= Cp15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CATTACNTAG TTNATTCTTG ACAGAATTAC GATTCGCTGG TATAATATAT CAGTACTGTT

60

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION:4..58

(D) OTHER INFORMATION:/standard\_name= "Constitutional  
promoter"  
/label= Cp16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

CATTGTGTAG TTTATTCTTG ACAGCTATGA GTCAATTGG TATAATAACA GTACTCAG

58

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:



(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION:4..59
- (D) OTHER INFORMATION:/standard\_name= "Constitutional promoter"  
/label= Cp17

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CATTCTGGAG TTTATTCTTG ACCGCTCAGT ATGCAGTGGT ATAATAGTAC AGTACTGTT

59

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION:4..58
- (D) OTHER INFORMATION:/standard\_name= "Constitutional promoter"  
/label= Cp18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CATTTTGCAG TTTATTCTTG ACATTGTGTG CTCGGGTGT ATAATACTAA GTACTGTT

58

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION:4..58

(D) OTHER INFORMATION:/standard\_name= "Constitutional  
promoter"  
/label= Cp19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CATCGCTTAG TTTTCTTGA CAGGAGGGAT CCGGGTTGAT ATAATAGTTA GTACTGTT

58

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 60 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION:4..60

(D) OTHER INFORMATION:/standard\_name= "Constitutional  
promoter"  
/label= Cp2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CATTGCTAG TTTATTCTTG ACATGAAGCG TGCCTAATGG TATATTACTT GAGTACTGTT

60

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 60 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION:4..60

(D) OTHER INFORMATION:/standard\_name= "Constitutional  
promoter"  
/label= Cp20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CATGGGTGAG TTTATTCTTG ACAGTGC GGC CNGGGGCTGA TATCATAGCA GAGTACTATT 60

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 4..59
- (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CATTACCGAG TTTATTCTTG ACACCGTTTA TCGGGGTTGT ATAATACTAT AGTACTGTT 59

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 4..60
- (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CATGTAGGAG TTTATTCTTG ACAGATTAGT TAGGGGGTGG TATAATATCT CAGTACTGTT 60

## (2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactococcus lactis
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION: 4..60
  - (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp24

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

CATGGGTAAG TTTATTCTTC ACACTATCTG GGCCCGATGG TATAATAAGT GACTACTGTT

60

## (2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 59 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactococcus lactis
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION: 3..59
  - (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

CTTTGGCAGT TTATTCTTGA CATGTAGTGA GGGGGCTGGT ATAATCACAT AGTACTGTT

59

## (2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Lactococcus lactis

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp26

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CATTCTACAG TTTATTCTTG ACATTGCACT GTCCCCCTGG TATAATAACT ATACATGCAT

60

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Lactococcus lactis

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CATGGGGCCG TTTATTCTTG ACAACGGCGA GCAGACCTGG TATAATAATA TAGTACTGTT

60

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..59

(D) OTHER INFORMATION: /standard\_name= "Constitutional  
promoter"  
/label= Cp29

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CATCGGTAAG TTATTCTTGA CATCTCAGGG GGGACGTGGT ATAATACTG AGTACTGTT

59

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional  
promoter"  
/label= Cp3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

CATCCTGTAG TTTATTCTTG ACACACGTNN TTAGCTGTGG TATAATAGGA GAGTACTGTT

60

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION:4..60

(D) OTHER INFORMATION:/standard\_name= "Constitutional  
promoter"  
/label= Cp30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CATGACAGAG TTTATTCTTG ACAGTATTGG GTTACTTTGG TATAATAGTT GAGTACTGTT 60

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION:4..60

(D) OTHER INFORMATION:/standard\_name= "Constitutional  
promoter"  
/label= Cp32

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

CATACGGGAG TTTATTCTTG ACATATTGCC GGTGTGTTGG TATAATAACT TAGTACTGTT 60

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION:4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional  
promoter"  
/label= Cp33

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

CATGTTGGAG TTTATTCTTG ACATACAATT ACTGCAGTGA TATAATAGGT GAGTACTGTT 60

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 60 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional  
promoter"  
/label= Cp34

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

CATCGCGAAG TTTATTCTTC ACACACCGCA GAACTTGTGG TATAATACAA CAGTACTGTT 60

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 59 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..59

(D) OTHER INFORMATION: /standard\_name= "Constitutional  
promoter"  
/label= Cp37



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CATCATTAAAG TTTATTCTTC ACATTGGCCG GAATTGTTGT ATAATACCTT AGTACTGTT

59

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 4..60
- (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp38

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CATAGAGAAG TTTATTCTTG ACAGCTAACT TGGCCTTGA TATAATACAT GAGTACTGTT

60

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 4..60
- (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp39

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

CATTGCGAAG TTTATTCTTG ACAGTACGTT TTTACCATGA TATAATAGTA TAGTACTGTT

60

## (2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Lactococcus lactis*
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION: 4..60
  - (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp4
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

GATGTTTTAG TTTATTCTTG ACACCGTATC GTGCGCGTGA TATAATCGGG ATCCTTAAGA 60

## (2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 59 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Lactococcus lactis*
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION: 4
  - (D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp40
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CATAGAACAG TTTATTCTTG ACATTGAATA AGAAGGCTGA TATAATAGCC AGTACTGTT 59

## (2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactococcus lactis
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION:4..60
  - (D) OTHER INFORMATION:/standard\_name= "Constitutional promoter"  
/label= Cp41
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

CATCCGCAAG TTTATTCTTG ACAGCTGAAT GTAGACGTGG TATAATAGTT AAGTACTGTT' 60

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactococcus lactis
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION:4..60
  - (D) OTHER INFORMATION:/standard\_name= "Constitutional promoter"  
/label= Cp42
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

CATTCGTAAG TTTATTCTTG ACACCTGAGA TGAGGCGTGA TATAATAAAT AAGTACTGTT 60

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 59 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..59

(D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp44

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

CATCGGGTAG TTTATTCTTG ACAATTAAGT AGAGCCTGAT ATAATAGTTC AGTACTGTT

59

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..59

(C) IDENTIFICATION METHOD: experimental

(D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Constitutional promoter"  
/label= Cp5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

CATGGGGGAG TTTATTCTTG ACATCATCTT CGTAGCCTGG TATACTACAT GAGTATGTT

59

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

## (ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp6

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

CATGTGGGAG TTTATTCTTG ACACAGATAT TTCCGGATGA TATAATAACT GAGTACTGTT

60

## (2) INFORMATION FOR SEQ ID NO: 40:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: YES

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

## (ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 4..60

(D) OTHER INFORMATION: /standard\_name= "Constitutional promoter"  
/label= Cp7

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

TATGCGGTAG TTTATTCTTG ACATGACGAG ACAGGTGTGG TATAATGGGT CTAGATTAGG

60

## (2) INFORMATION FOR SEQ ID NO: 41:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: YES

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus lactis*

## (ix) FEATURE:

(A) NAME/KEY: promoter

- (B) LOCATION:4..60
- (D) OTHER INFORMATION:/standard\_name= "Constitutional promoter"  
/label= Cp8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

CATTCTTTAG TTTATTCTTG ACAAACGTAT TGAGGACTGA TATAATAGGT GAGTACTGTT

60

(2) INFORMATION FOR SEQ ID NO: 42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 60 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Lactococcus lactis
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION:4..60
  - (D) OTHER INFORMATION:/standard\_name= "Constitutional promoter"  
/label= Cp9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

CATAGTCTAG TTTATTCTTG ACACGCGGTC CATTGGCTGG TATAATAATT TAGTACTGTT

60

(2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 177 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Saccharomyces cerevisiae
- (ix) FEATURE:
  - (A) NAME/KEY: promoter
  - (B) LOCATION:8..177
  - (C) IDENTIFICATION METHOD: experimental
  - (D) OTHER INFORMATION:/evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp102

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

GAATTCGTGA CTCAAACGGG TGGTCGACGG GTGGTTCCAA TTAATTGGCG TCCCTCTTAT 60  
AAAGGCGAGG GTACGTGCGA CAATTGGTAG AGCGAGCGGG GCTCTTAAGT GCAAGTGACT 120  
GCGAACATTT TTTTCGTTTG TTAGAATAAT TCAAGAATCG CTACCAATCA TGGATCC 177

## (2) INFORMATION FOR SEQ ID NO: 44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saccharomyces cerevisiae*

## (ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..181
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp112

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

GAATTCGTGA CTCACGGCAT CTGATGGTTG ACCATAGTCA GGAACATTGT GCTGGAGTTC 60  
CTTGAGGAAT GAGTTATAAA ATGGGAGGTT GCGGCTAATG CCAGGCAGGA GAGGAACCCT 120  
CTTAAGTGCA AGTGACTGCA AACATTTTTT TCGTTTGTG AATCGCTACC AATCATGGAT 180  
CC 182

## (2) INFORMATION FOR SEQ ID NO: 45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 191 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saccharomyces cerevisiae*

## (ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..181
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

```
GAATTCGTGA CTCACTAGGC AGGTCACGTT GGCTCTTCGC GCGCAGGTT CGTATGCCGC      60
GCCGCCAGGG GCTTTATAAA GGTCGTCCTG GGTACAGTTG GGATGGCTCC ACGTTTCGGC      120
TCTTAAGTGC AAGTGACTGC GAACATTTTCG TTTGTTAGAA TAATTCAAGA ATCGCTACCA      180
ATCATGGATC C                                                                191
```

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 167 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..167
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

```
GAATTCGTGA CTCAGGGCCG TACTAAGTAG CTTTCGTATG CTATGCGGGG TTTTATAAAT      60
CTTTGGGCCA TGGTCTTGCT GGAAAACACC TCTCTTAAGT GCAAGTGACT GCGAACATTT      120
TTTTCGTTTG TTAGAATAAT TCAAGAATCG CTACCAATCA TGGATCC                      167
```

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 191 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES



(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 8..191

(C) IDENTIFICATION METHOD: experimental

(D) OTHER INFORMATION: /evidence= EXPERIMENTAL

/standard\_name= "Yeast promoter"

/label= Yp154

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

GAATTCGTGA CTCACCGCTC GGGTGCAGGG CCAAGGCGGC GGAATGTGCG GGGCGTTCTA	60
GCGCAATCGG GGTATAAATT TATAAGGAGG CTGCGGGTGC TAGTTTGTCT AGTTTGACTC	120
TTAAGTGCAA GTGACTGCGA ACATTTTTCG TTTGTTAGAA TAATTCAAGA ATCGCTACCA	180
ATCATGGATC C	191

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 195 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

(A) NAME/KEY: promoter

(B) LOCATION: 8..190

(C) IDENTIFICATION METHOD: experimental

(D) OTHER INFORMATION: /evidence= EXPERIMENTAL

/standard\_name= "Yeast promoter"

/label= Yp18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

GAATTCGTGA CTCAGGATTA GCTATGCCGG TTGGGATAAG CGAACAACTG GAGGTGAGAA	60
GCTTTTGTGC AGAATATAAA CCCGTTAGTC AGGGTTTGGT GGGATAGGGG GTACTGTACC	120
TCTTAAGTGC AAGTGACTGC GAACATTTTT TTCGTTTGTG AGAATAATTC AAGAATCGCT	180
ACCAATCATG GATCC	195

(2) INFORMATION FOR SEQ ID NO: 49:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 179 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: double  
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
    (A) ORGANISM: *Saccharomyces cerevisiae*
- (ix) FEATURE:  
    (A) NAME/KEY: promoter  
    (B) LOCATION: 8..179  
    (C) IDENTIFICATION METHOD: experimental  
    (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
        /standard\_name= "Yeast promoter"  
        /label= Yp183
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

GAATTCGTGA CTCACTAAGG GTTCGCCATT AACAGAATCG CTGGTAGAAC ATCGGTAGTT	60
AGGCACCCGA GTATAAACAG GCGGACCCCT CACGGATATC AGCTGATAGT GCGAGCCTCA	120
ATGCGAACAT TTTTTCGTT TGTTAGAATA ATTCAAGAAT CGCTACCAAT CATGGATCC	179

(2) INFORMATION FOR SEQ ID NO: 50:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 195 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: double  
    (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
    (A) ORGANISM: *Saccharomyces cerevisiae*
- (ix) FEATURE:  
    (A) NAME/KEY: promoter  
    (B) LOCATION: 8..190  
    (C) IDENTIFICATION METHOD: experimental  
    (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
        /standard\_name= "Yeast promoter"  
        /label= Yp190
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

GAATTCGTGA CTCAGTATCC ACGGGTGTTC GAGGGCTGGT CGCAGGTTAG CAGGCGAGGG	60
CGGGTGGTTA CGGCTATAAA TGAGTGTTTG CAGCCGGGTA CGGGCGTACG AGTAGTGATC	120

TCTTAAATGC AAGTGACTGC GAACATTTTT TTCGTTTGTT AGAATAATTC AAGAATCGCT 180  
ACCAATCATG GATCC 195

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 193 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..189
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yest promoter"  
/label= Yp191

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

GAATTCGTGA CTCAATGCTG CGGGCGGCAG GAGTCTGGTG TAACTTCCCA TTTTGAGTGA 60  
AAGACAGACC ATCTATAAAC ATTTGGTGGG CAAAGTGGCC TGGCGGATTT GTTTGGACTC 120  
TTAAGTGAAA GTGACTGCGA ACATTTTTTT CGTTTGTTAG AATAATTCAA GAATCGCTAC 180  
CAATCATGGA TCC 193

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..166
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL

/standard\_name= "Yeast promoter"  
/label= Yp192

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

```
GAATTCGTGA CTCACTTAAG GCTACTGCGG AAGTTTAGAT CTAAGGTCGG AAATAATTTA    60
GAAAATTACG ACATTATAAA TAGCGGAGAG GCCAGGTGAT GGGCACCATT GTGGGGGGGC    120
TCTTAATTGT TAGAATAATT CAAGAATCGC TACCAATCAT GGATCC                    166
```

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..190
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp212

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

```
GAATTCGTGA CTCAGTCGCC CGCAAGATGG GATGGTGCAT TTAAACACC CGAATTATAC    60
TCGTCAACTT ATAGTATAAA CGGAACGCGA CGATACGTTT TAGTTTTTCGG CGAAGTCGAC    120
TCTTAAGTGC AAGTGACTGC GAACATTTTT TTCGTTTGTT AGAATAATTC AAGAATCGCT    180
ACCAATCATG GATCC                                                    195
```

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 188 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..183
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

```
GAATTCGTAC TCACGACAGC GTTATGACTT CGAGGACCAG CTA CTTCGG TCGCGTACTA      60
GTTTTTACCT GTATAACTT TGCTACCGCT GGGCCTTGGT GGTGCTGTCC CGCTCTTAAG      120
TGCAAGTGAC TGC GAACATT TTTTCGTTT GTTACAATAA TTCAAGAATC GCTACCAATC      180
ATGGATCC                                188
```

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

(ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..190
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp435

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

```
GAATTCGTGA CTAAATGGAT AAGTTATCG CCATCACGGA GTCTTCTCTC ACGTCTGGAG      60
CAGAGGCTAG ACCTTATAAA TTATACATGG TGGGAGAGGC GATAGTCTTT AGAGACGTGC      120
TCTTAAGTGC AAGTGACTGC GAACATTTTT TTCGTTTGT AGAATAATTC AAGAATCGCT      180
ACCAATCATG GATCC                                195
```

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 189 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Saccharomyces cerevisiae*
- (ix) FEATURE:  
(A) NAME/KEY: promoter  
(B) LOCATION: 8..184  
(C) IDENTIFICATION METHOD: experimental  
(D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp68

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

GAATTCGTGA CTCACAAGAA TGTGGGCGGG TCGTTAACT GAGCCTGGAC ACCTTGGCTG	60
CGTCGCTTTC GTATAAGAT CTTAGAGCTG TGGAGTCTGG GTCGAGTGGC CAGCTCTTAA	120
ATGCAAGTGA CTGCGAACAT TTTTTCGTT TGTTAGAATA ATTCAAGAAT CGCTACCAAT	180
CATGGATCC	189

(2) INFORMATION FOR SEQ ID NO: 57:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 195 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: YES
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Saccharomyces cerevisiae*
- (ix) FEATURE:  
(A) NAME/KEY: promoter  
(B) LOCATION: 8..190  
(C) IDENTIFICATION METHOD: experimental  
(D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp89

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

GAATTCGTGA CTCACTCGGA AGATTGGGTT TACGATTAGG ATGGCGCGGC AGAACCGGGG	60
GGGATTCCCT TCTATATAAA GGGTCCGAT ACTACGTGCT GCGGACGGCC GATCGAGTTA	120
TCTTAAGTGC AAGTGACTGC GAAAATTTT TTCGTTTGTT AGAATAATTC AAGAATCGCT	180

ACCAATCATG GATCC

195

## (2) INFORMATION FOR SEQ ID NO: 58:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 176 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

## (ix) FEATURE:

- (A) NAME/KEY: promoter
- (B) LOCATION: 8..171
- (C) IDENTIFICATION METHOD: experimental
- (D) OTHER INFORMATION: /evidence= EXPERIMENTAL  
/standard\_name= "Yeast promoter"  
/label= Yp96

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

GAATTCGTGA CTCATCTAGT GAGAGGAGCC GTGGTATCTT GTGTCACCAC CAGGGGAAAA	60
TAATGGCAGG GGTGTATAAA TGGTCGAGTA GTCGCGACCC ACGCTGCAAG GCAAGGAACT	120
CTTAAATTTT TTTCGTTTGT TAGAATAATT CAAGAATCGC TACCAATCAT GGATCC	176

## PATENT CLAIMS

1. An artificial promoter library for a selected organism or group of organisms, comprising a mixture of double stranded DNA fragments the sense strands of which comprise at least two consensus sequences of efficient promoters from said organism or group of organisms, or parts thereof comprising at least half of each, and surrounding or intermediate nucleotide sequences (spacers) of variable length in which at least 7 nucleotides are selected randomly among the nucleobases A, T, C and G, with the proviso that previously known promoter sequences and promoter sequences isolated from natural sources are not comprised.
2. An artificial promoter library according to claim 1, wherein at least 10, preferably at least 12, and more preferably at least 14, nucleotides in the spacer sequence(s) are selected randomly among the nucleobases A, T, C and G.
3. An artificial promoter library according to claim 1 or 2, wherein the sense strands of the double stranded DNA fragments include a regulatory DNA sequence imparting a specific regulatory feature to the promoters of the library.
4. An artificial promoter library according to claim 3, wherein the specific regulatory feature is activation by a change in the growth conditions.
5. An artificial promoter library according to any one of claims 1-4, wherein the double stranded DNA fragments have a sequence comprising one or more recognition sites for restriction endonucleases added to one of or both their ends.



6. An artificial promoter library according to claim 5, wherein the double stranded DNA fragments have a sequence specifying multiple recognition sites for restriction endonucleases (multiple cloning sites MCS) added to one of  
5 or both their ends.

7. An artificial promoter library according to any one of claims 1-6, wherein the selected organism or group of organisms is selected from the group consisting of pro-  
10 karyotes.

8. An artificial promoter library according to claim 7, wherein the prokaryotic microorganism is selected from the group consisting of bacteria belonging to the genera  
15 *Lactococcus*, *Streptococcus*, *Enterococcus*, *Lactobacillus*, *Leuconostoc*, *Escherichia*, *Bacillus* and *Pseudomonas*.

9. An artificial promoter library according to claim 8, wherein the prokaryotic microorganism is selected from  
20 the group consisting of *Lactococcus lactis*, *Streptococcus thermophilus*, *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas putida*.

10. An artificial promoter library according to any one  
25 of claims 7-9, wherein said consensus sequences comprise the -10 signal (-12 to -7): TATAAT and at least one activator protein binding site upstream of the -10 signal or parts thereof comprising at least 3 conserved nucleotides of each.

30  
11. An artificial promoter library according to any one of claims 7-9, wherein said consensus sequences comprise the -35 signal (-35 to -30): TTGACA and the -10 signal (-12 to -7): TATAAT or parts of both comprising at least  
35 3 conserved nucleotides of each.

12. An artificial promoter library according to claim 11, wherein said consensus sequences comprise from 4 to 6 conserved nucleotides of the -35 signal and from 4 to 6 conserved nucleotides of the -10 signal, preferably 5 or 6, and more preferably all 6 nucleotides of each.
13. An artificial promoter library according to any one of claims 10-12, wherein said consensus sequences further comprise intervening conserved motifs.
14. An artificial promoter library according to claim 13, wherein said intervening conserved motifs are selected from the group consisting of the conserved motifs -44 to -41: AGTT, -40 to -36: TATTC, -15 to -14: TG, and +1 to +8: GTACTGTT.
15. An artificial promoter library according to any one of claims 11-14, wherein the length of the spacer between the -35 signal and the -10 signal is 14-23 bp, preferably 16-18 bp, and more preferably 17 bp.
16. An artificial promoter library according to any one of claims 11-15, wherein the sense strands of the double stranded DNA fragments have the degenerated sequence shown in SEQ ID No. 1 with minor variations in the consensus sequences and spacer lengths.
17. An artificial promoter library according to any one of claims 11-15, wherein the sense strands of the double stranded DNA fragments, including an upstream temperature regulatory DNA sequence, have the degenerated sequence shown in SEQ ID No. 2 with minor variations in the consensus sequences and spacer lengths.
18. An artificial promoter library according to any one of claims 1-6, wherein the selected organism or group of

organisms is selected from the group consisting of eukaryotes.

19. An artificial promoter library according to claim 18,  
5 wherein the eukaryotic organism or group of organisms is selected from the group consisting of eukaryotic microorganisms.

20. An artificial promoter library according to claim  
10 19, wherein the eukaryotic microorganism or group of microorganisms is selected from the group consisting of yeasts, other fungi and mammalian cell lines.

21. An artificial promoter library according to claim  
15 20, wherein the eukaryotic microorganism is a yeast of the species *Saccharomyces cerevisiae*.

22. An artificial promoter library according to any one  
20 of claims 18-21, wherein said consensus sequences comprise a TATA box and at least one upstream activation sequence (UAS).

23. An artificial promoter library according to claim  
25 22, wherein said consensus sequences further comprise a transcription initiation signal (TI box) functioning in *Saccharomyces cerevisiae*.

24. An artificial promoter library according to claim  
30 23, wherein said consensus sequences comprise the TATA box: TATAAA, the UAS<sub>GCN4p</sub>: TGACTCA, and the TI box: CTCTTAAGTGCAAGTGACTGCGA, which also functions as the binding site for the arginine repressor, *argR*.

25. An artificial promoter library according to any one  
35 of claims 22-24, wherein the spacer length between the TATA box and the first UAS and between this and further

UAS, if any, is up to 1000 bp, preferably from 10 to 500 bp, and more preferably from 15 to 150 bp.

26. An artificial promoter library according to any one  
5 of claims 22-25, wherein the sense strands of the double stranded DNA fragments have the degenerated sequence shown in SEQ ID No. 3 with minor variations in the consensus sequences and spacer lengths.

10 27. An artificial promoter selected from an artificial promoter library according to any one of claims 1-26.

28. A promoter according to claim 27, comprising a DNA  
15 sequence selected from the group consisting of the SEQ IDs Nos. 5-58.

29. An artificial promoter derived from an artificial promoter according to claim 27 or 28.

20 30. A method of constructing an artificial promoter library for a selected organism or group of organisms, which comprises selecting at least two consensus sequences of efficient promoters from said organism or group of organisms; synthesizing a mixture of single  
25 stranded DNA sequences comprising said consensus sequences, or parts thereof comprising at least half of each, and surrounding or intermediate nucleotide sequences (spacers) of variable length in which at least 7 nucleotides are selected randomly among the nucleobases  
30 A, T, C and G; and converting the single stranded DNA sequences into double stranded DNA fragments by second strand synthesis.

31. A method according to claim 30, wherein at least 10,  
35 preferably at least 12, and more preferably at least 14, nucleotides in the spacer sequence(s) are selected randomly among the nucleobases A, T, C and G.

32. A method according to claim 30 or 31, wherein the single stranded DNA sequences include a regulatory DNA sequence imparting a specific regulatory feature to the promoters of the library.

5

33. A method according to claim 32, wherein the specific regulatory feature is activation by a change in the growth conditions

10 34. A method according to any one of claims 30-34, wherein a sequence specifying one or more recognition sites for restriction endonucleases is added to one of or both the ends of the single stranded DNA sequences in the synthesis, or a linker comprising such restriction sites  
15 is ligated to one of or both the ends of the double stranded DNA fragments.

35. A method according to claim 34, wherein a sequence specifying multiple recognition sites for restriction en-  
20 donucleases is added to one of or both the ends of the single stranded DNA sequences in the synthesis, or a linker comprising such multiple cloning sites (MCS) is ligated to one of or both the ends of the double stranded DNA fragments.

25

36. A method according to any one of claims 30-35, wherein the selected organism or group of organisms is selected from the group consisting of prokaryotes.

30 37. A method according to claim 36, wherein the prokary-  
otic microorganism or group of microorganisms is selected from the group consisting of bacteria belonging to the genera *Lactococcus*, *Escherichia*, *Bacillus* and *Pseudo-*  
*monas*.

35

38. A method according to claim 37, wherein the prokary-  
otic microorganism is selected from the group consisting

of *Lactococcus lactis*, *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas putida*.

39. An artificial promoter library according to any one  
5 of claims 36-38, wherein said consensus sequences comprise the -10 signal (-12 to -7): TATAAT and at least one activator protein binding site upstream of the -10 signal or parts thereof comprising at least 3 conserved nucleotides of each.

10

40. A method according to any one of claims 36-38,  
wherein said consensus sequences comprise the -35 signal  
(-35 to -30): TTGACA and the -10 signal (-12 to -7):  
TATAAT or parts of both comprising at least 3 conserved  
15 nucleotides of each.

41. A method according to claim 40, wherein said consensus sequences comprise from 4 to 6 conserved nucleotides of the -35 signal and from 4 to 6 conserved nucleotides  
20 of the -10 signal, preferably 5 or 6, and more preferably all 6 nucleotides of each.

42. A method according to any one of claims 39-41,  
wherein said consensus sequences further comprise intervening conserved motifs.  
25

43. A method according to claim 42, wherein said intervening conserved motifs are selected from the group consisting of the conserved motifs -44 to -41: AGTT, -40 to  
30 -36: TATTC, -15 to -14: TG, and +1 to +8: GTACTGTT.

44. A method according to any one of claims 40-43,  
wherein the length of the spacer between the -35 signal and the -10 signal is 14-23 bp, preferably 16-18 bp, and  
35 more preferably 17 bp.

45. A method according to any one of claims 40-44, wherein the mixture of single stranded DNA sequences has the degenerated sequence shown in SEQ ID No. 1 with minor variations in the consensus sequences and spacer lengths.

5

46. A method according to any one of claims 40-42 and 44, wherein the mixture of single stranded DNA sequences including an upstream temperature regulatory DNA sequence has the degenerated sequence shown in SEQ ID No. 2 with  
10 minor variations in the consensus sequences and spacer lengths.

47. A method according to any one of claims 30-35, wherein the selected organism or group of organisms is  
15 selected from the group consisting of eukaryotes.

48. A method according to claim 47, wherein the eukaryotic organism or group of organisms is selected from the group consisting of eukaryotic microorganisms.

20

49. A method according to claim 48, wherein the eukaryotic microorganism or group of microorganisms is selected from the group consisting of yeasts, other fungi and mammalian cell lines.

25

50. A method according to claim 49, wherein the eukaryotic microorganism is a yeast of the species *Saccharomyces cerevisiae*.

30 51. A method according to any one of claims 47-50, wherein said consensus sequences comprise a TATA box and at least one upstream activation sequence (UAS).

52. A method according to claim 51, wherein said consensus sequences further comprise a transcription initiation  
35 signal (TI box) functioning in *Saccharomyces cerevisiae*.

53. A method according to claim 52, wherein said consensus sequences comprise

the TATA box: TATAAA, the UASGCN4p: TGA CTCA, and the TI box: CTCTTAAGTGCAAGTGACTGCGA, which also functions as the binding site for the arginine repressor, *argR*.

54. A method according to any one of claims 51-53, wherein the spacer length between the TATA box and the first UAS and between this and further UAS, if any, is up to 1000 bp, preferably from 10 to 500 bp, and more preferably from 15 to 150 bp.

55. A method according to any one of claims 51-54, wherein the sense strands of the double stranded DNA fragments have the degenerated sequence shown in SEQ ID No. 3 with minor variations in the consensus sequences and spacer lengths.

56. A method of optimizing the expression of a gene in a microorganism, which comprises

a) selecting a set of promoters covering a range of promoter activities in relatively small steps of activity change from an artificial promoter library according to any one of claims 1-26 or constructed by the method according to any one of claims 30-55;

b) cloning said set of promoters into said organism placing in each clone said gene under the control of at least one promoter from the set;

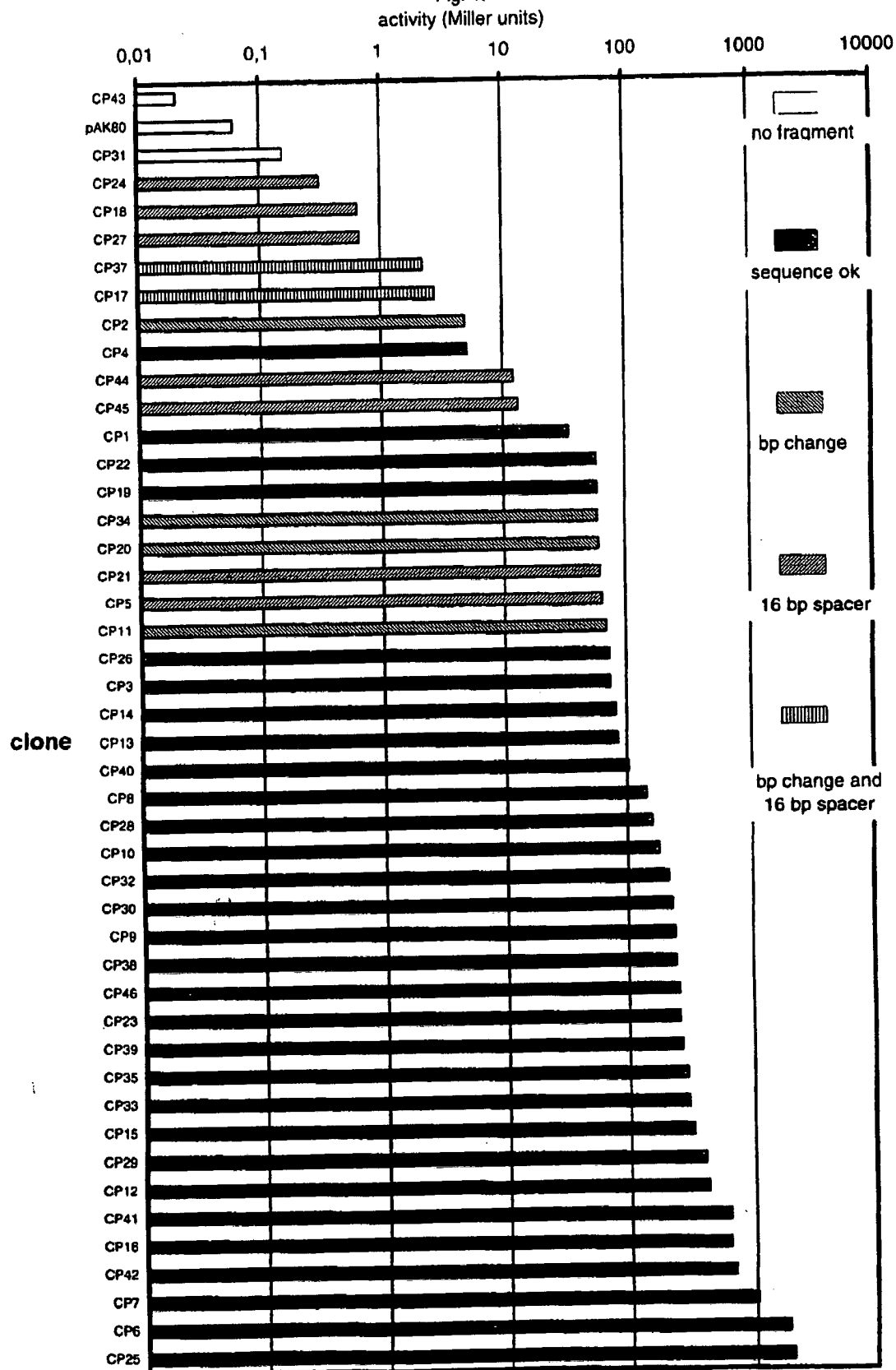
c) growing the selected clones and screening them to find the one showing optimized flux of product formation.



57. A method according to claim 56, wherein the selected organism is selected from the group consisting of prokaryotic and eukaryotic microorganisms.

1/5

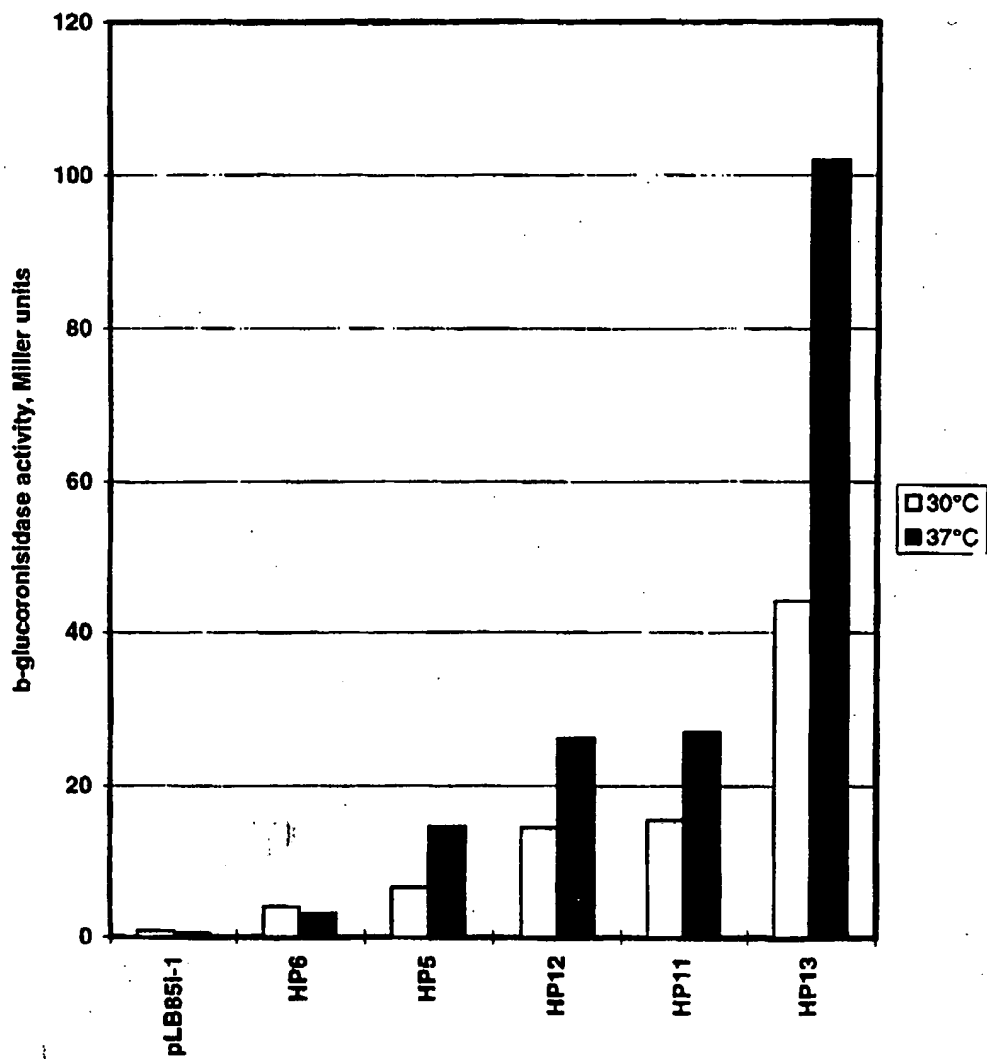
Fig. 1.



SUBSTITUTE SHEET

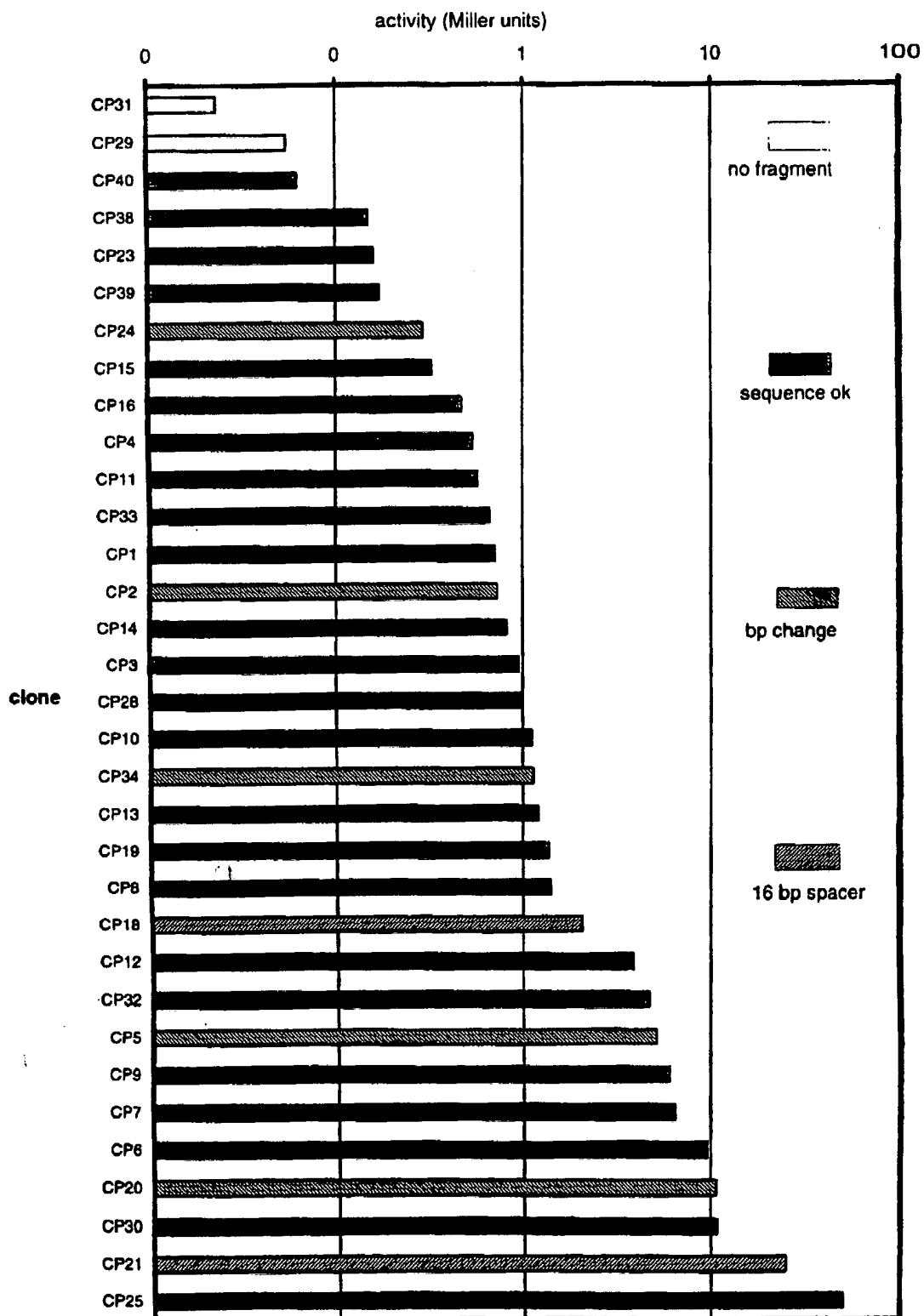
2/5

FIG. 2

Heatshock regulated promoters in *L. lactis*

3/5

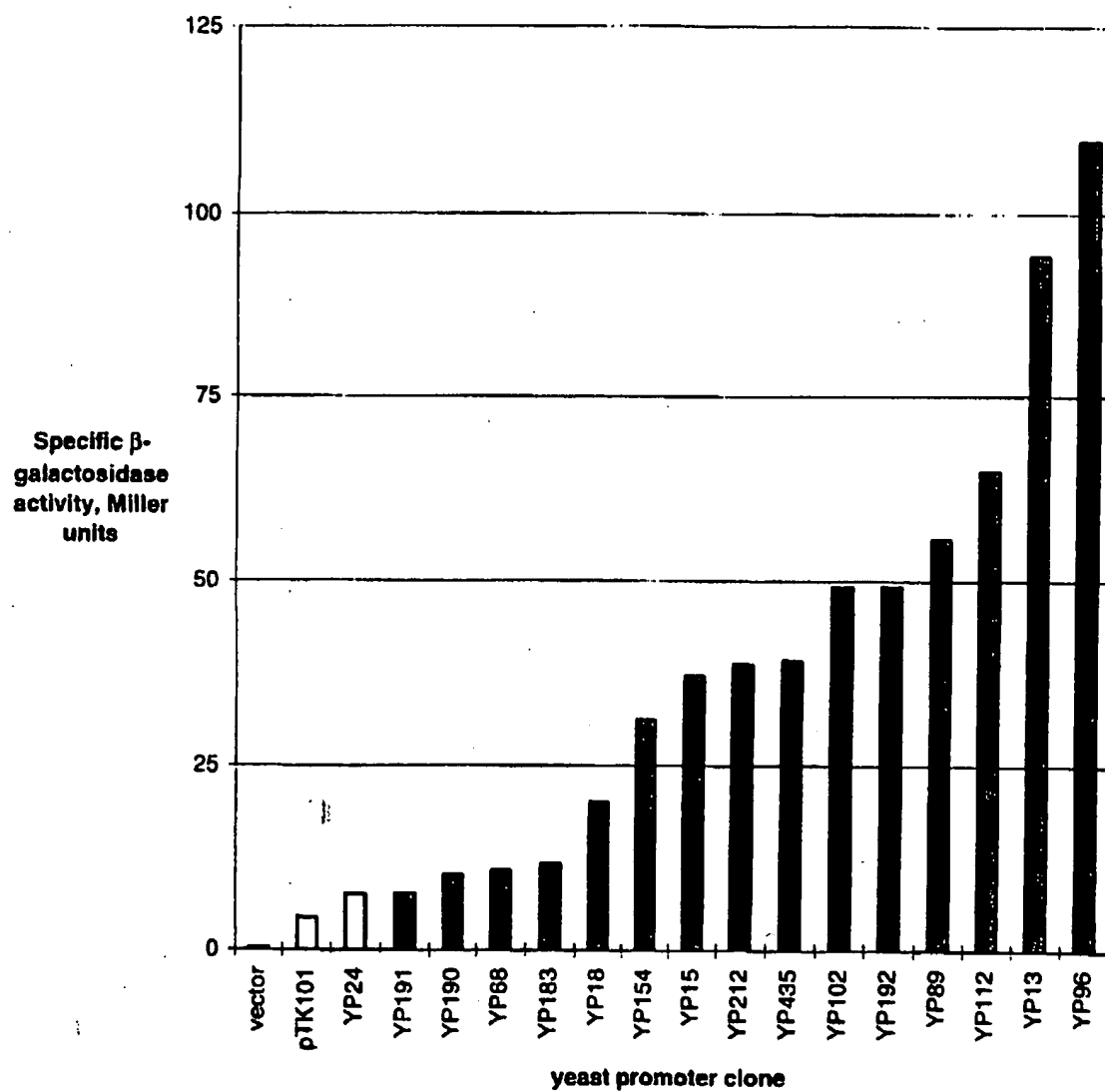
Fig. 3.



SUBSTITUTE SHEET

Fig. 4

## Artificial yeast promoters



5/5

FIG. 5

Artificial yeast promoters, regulation by  
external arginine

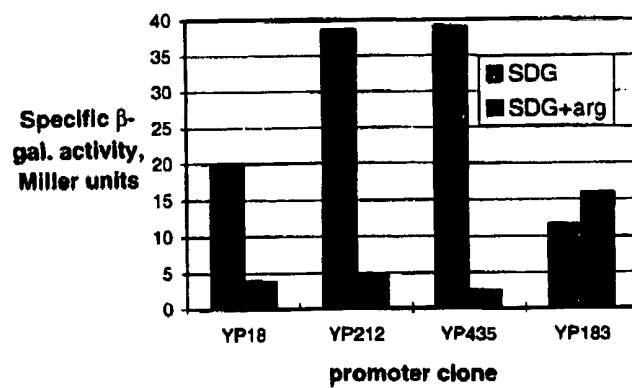
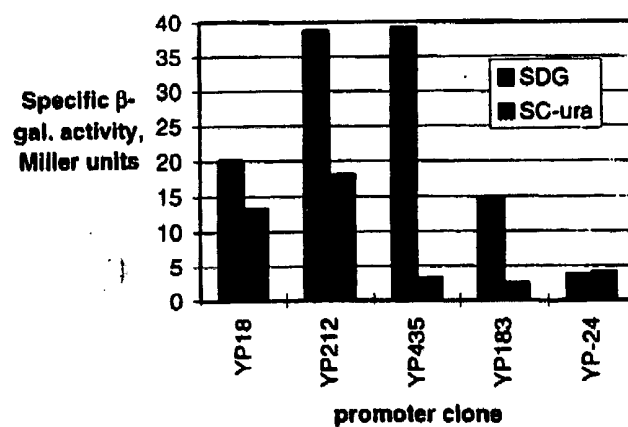


FIG. 6

Artificial yeast promoters, general regulation  
by external amino acids



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 97/00342

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
IPC6: C12N 15/11 // C12N 15/63 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
IPC6: C12N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
SE,DK,FI,NO classes as above		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPI, EPODOC, MEDLINE, EMBASE, BIOSIS, SCISEARCH		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 9425609 A1 (HYBRITECH INCORPORATED), 10 November 1994 (10.11.94)  --	1-57
X	Nucleic Acids Research, Volume 16, No 15, 1988, Arnold R. Oliphant et al, "Defining the consensus sequences of E.coli promoter elements by random selection" page 7673 - page 7683  --	1-57
A	Dialog Information Services, file 155, MEDLINE, Dialog accession no. 08012585, Medline accession no. 94368841, Nilsson D et al: "A conserved sequence in tRNA and rRNA promoters of Lactococcus latis"; & Biochim Biophys Acta Sep 13 1994, 1219 (1) p141-4  --	16
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" documents of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
2 February 1998		04 -02- 1998
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86		Authorized officer  Carl-Olof Gustafsson Telephone No. +46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/DK 97/00342

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>Dialog Information Services, file 155, MEDLINE, Dialog accession no. 07764569, Medline accession no. 94172317, Eaton T et al: "Cloning and sequence analysis of the dnak gene region of Lactococcus lactis subsp. lactis"; &amp; J Gen Microbiol Dec 1993, 139 (Pt 12) p3253-64</p> <p>--</p>	17
A	<p>Dialog Information Services, file 155, MEDLINE, Dialog accession no. 08484065, Medline accession no. 96076630, Casas C et al: "Sequence analysis of a 9873 bp fragment of the left arm of yeast chromosome XV that contains the ARG8 and CDC33 genes, a putative riboflavin synthase beta chain gene, and four new open reading frames"; &amp; Yeast Sep 15 1995, 11 (11) p1061-7</p> <p>-- -----</p>	26



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/ DK 97/00342

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-57

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

International application No.  
PCT/DK 97/00342

Form PCT/ISA/210 (patent family annex) (July 1992)

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☒ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**BLANK PAGE**